SEARCH FOR EXTRATERRESTRIAL LIFE USING CHIRAL MOLECULES:

MANDELATE RACEMASE AS A TEST CASE

A Thesis
Presented to
The Academic Faculty

by

Tracey Lyn Thaler

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy in the
School of Chemistry & Biochemistry

Georgia Institute of Technology
May 2007

SEARCH FOR EXTRATERRESTRIAL LIFE USING CHIRAL MOLECULES:

MANDELATE RACEMASE AS A TEST CASE

Approved by:

Dr. Andreas S. Bommarius, Advisor
School of Chemical & Biomolecular Engineering
Georgia Institute of Technology

Dr. Christoph J. Fahrni
School of Chemistry & Biochemistry
Georgia Institute of Technology

Dr. Rick P. Trebino
School of Physics
Georgia Institute of Technology

Dr. Phillip R. Gibbs
Chief Science Officer
Stheno Corporation

Dr. Donald F. Doyle
School of Chemistry & Biochemistry
Georgia Institute of Technology

Date Approved: 03-28-07

ACKNOWLEDGEMENTS

First and foremost I would like to thank my advisor Andy Bommarius for his support and encouragement over the past six and half years. I am also thankful to my thesis committee members: Donald Doyle, Christoph Fahrni, Phillip Gibbs and Rick Trebino for their patience as well as guidance. I would especially like to the thank Phillip Gibbs, who was not only the mastermind behind my thesis topic, but has always been encouraging in suggesting ideas to enhance my work.

I would also like to thank my lab members past and present. Karen Polizzi has been a wonderful friend and colleague. We had many great conversations about α/β-barrel enzymes, which led to enhance my knowledge in this area. James Broering, also a great friend, has been very supportive of my efforts and helpful in teaching me about DSC, also I have found his cynical views refreshing. Javier Chaparro-Riggers was a great help in the development of the fourth Aim of this thesis as well as encouraging throughout the past two years. I would like to thank Janna Blum and Thomas Rogers for proofreading parts of my thesis. Also, thanks to Rongrong Jiang, Eduardo Vazquez and Bernard Loo, whom I have enjoyed working along side of in the lab.

I have also had support from members of the Doyle lab. Especially Bahareh Azizi who has been encouraging in more ways than she will ever

know, she is an awesome person and I hope she realizes the great potential of her future. Also, I had some great times with Lauren Schwimmer, and I am thankful to her for helping me when I had chemistry questions. Terry Watt has been an extremely useful source for discussing analysis of irreversible DSC data, he is also always willing to help when I have a question. Kenyetta Johnson and Priyanka Rohatgi have been great friends.

I am very thankful to Christian Uehara who wrote the computer code for the DSC data analysis. I appreciate all of the hard work he spent on it. I am also thankful for the Astrobiology meetings with Rick Trebino and Neeraj Kothari, this project has increased my scientific knowledge in the field. I would also like to thank Brian Lynch from the School of Biology at Georgia Tech for always being able to provide me with adequate sequences. I would like to also thank Perry Mars for sequencing at the FAME sequencing center (Emory University).

Further I would like to thank my family for their support, most of all my sister Lisa Thaler. She has been the most supportive person in my life throughout this entire endeavor and I am very lucky to have her in my life. I am most thankful for Lisa's help with proofreading my thesis. I would also like to thank Donald Murphy for his technological support with my computer as well as his friendship. Margaret and Hadley Wellborn have been extremely supportive throughout the past two years and I am

especially thankful for their help with Lily at all times. William Wellborn has been an astonishing addition to my life and I am grateful for all that he has done for me both in and out of the lab. He is a wonderful person and I am lucky to have him in my life.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

SUMMARY

The possible existence of extraterrestrial life forms has been of interest to humans for many millennia. In the past few decades space travel has provided an opportunity to search life outside of Earth. Chiral molecules are critical molecules in Earth-based life and are among the first chemical molecules sought after as proof of potential extraterrestrial life; however, identification of these chiral molecules is difficult due the lack of sensitive instruments. The objective of this work is to develop a benchmark reaction to be used as a guide in the development of instrumentation, such as a polarimeter, to be used in the search for extraterrestrial life. To achieve this objective, to investigate the enzyme mandelate racemase (MR), which catalyzes the racemization between the enantiomers of mandelate. MR is a member of the enolase superfamily, which contains a $(\alpha/\beta)_7$-β barrel domain, the fold most frequently found among all known protein structures.

Activity of the enzyme was measured at low temperatures and in non-aqueous media, as these are the conditions that represent extraterrestrial terrain. We find that mandelate racemase (MR) is active in concentrated ammonium salt solutions and water-in-oil microemulsions in a temperature range between –30°C to 70°C; however, the enzyme is not active in several organic cryosolvents. The stability of the structure of MR

was also explored. Using differential scanning calorimetry (DSC) we observe the unfolding of the enzyme was irreversible and therefore kinetically controlled. We also found proof for divergent evolution of the enolase superfamily, providing evidence for divergent evolution across the MR and muconate lactonizing enzyme (MLE) subfamilies has been demonstrated. However, we also conclude that reactions yielding a polarimetric signal, such as racemizations employed in this work, are suitable as a tool to find signs of life.

CHAPTER I

INTRODUCTION: DEVELOPING A BENCHMARK IN THE SEARCH FOR EXTRATERRESTRIAL LIFE

Man has imagined the possibility of the existence of extraterrestrial life since ancient times. Over the millennia, the debate has progressed from the strictly philosophical to the empirical sciences where we find it firmly rooted today. The Greek philosopher Thales argued in the 7th century B.C. that an infinite universe must include infinite populated worlds (*1*). In contrast, Aristotle's geocentric theory in the 4th century B.C. led to the alternate conclusion that the life found on Earth is a singular occurance in the universe. The Aristotelean view prevailed for many centuries. It was not until the invention of the telescope in the 16th century A.D., that scientists reconsidered the possibility of extraterrestrial life. In this work, we continue this quest through study of the change in chirality of homochiral molecules, using a racemase, as an experimental benchmark that can lead us toward more conclusive tests for extraterrestrial life.

Attempting to determine a clear definition of 'life' from a scientific perspective was an early and important hurdle that scientists faced as

they began to seriously consider the possibility of extraterrestrial life. As new types of molecules have been discovered, most notably viruses, a comprehensive definition of life has become even more complicated. The National Aeronautics and Space Administration (NASA) has developed the Origins program to investigate the existence of extraterrestrial environments. The NASA Origins program defines living systems as those capable of "chemical Darwinian evolution" involving self-sustaining chemical systems that undergo evolution at the molecular level (*2*). This definition is limited in that life on Earth probably resulted from physical and chemical "contingencies" present at the time of its origin on the planet (*3*). Hence, there is no single definition of 'life' which is widely accepted by the scientific community. Fortunately, evidence of extraterrestrial life fitting any modern definition of life would be of interest to the scientific community as well as the general public. Instead of developing another definition for 'life,' we will instead draw on the existing scientific literature and definitions spanning physiological, metabolic, biochemical, genetic, and thermodynamic characteristics, among others (*4-9*).

The field of astrobiology has arisen to address the complex challenge of developing this theoretical basis for the study of extraterrestrial life, and further to use this understanding to identify those locations in the solar system most likely to harbor life. The work in this field is

largely theoretical, or relies on inferences drawn from results in other fields of science because of the huge expense of performing direct experiments on planets or moons in our solar system to confirm the current or past existence of life in those locations. Understanding how terrestrial life originated provides clues in the search for extraterrestrial life; for example, astrobiologists believe that the regions to be included in the search for life should contain a water source and energy (redox), as well as biological polymers to carry out central biological functions (*3, 10*). Using these criteria, the bodies that are believed to be most like Earth are Mars, Saturn's icy moons Titan and Enceladus, as well as Jupitor's moon, Europa. Mars is thought to have once held similar conditions to those of early Earth, with evidence of water discovered by the Mars exploration land rovers, Spirit and Opportunity (*11*)

Chiral molecules have been found to be critical to all Earth based life (*12*). Since Watson and Crick discovered the structure of DNA (*13, 14*) many advances have been made adding to our understanding of the chemical composition of living organisms. For example, DNA molecules include only D-deoxyribose sugars, while proteins are composed of L-amino acids. Chirality in biological systems is also vital in that exposing living organisms to interact with their complementary enantiomers often has destructive effects (*12*). Theories for the origin of homochirality are numerous in the literature and will be further discussed in Chapter 2. The

importance of these types of compounds has been understood for some time and the capacity to perform experiments to look for them is included in space missions (*2, 10, 15-18*).

Although attempts to detect organic molecules on Mars were inconclusive, trace amounts of L-alanine, DL-serine, and glycine were reportedly found on a Martian meteorite ALH84001 (*19*); however, it has since been determined that these findings are terrestrial contaminants (*15*). Further analysis of the Murchison meteorite, from which samples were retrieved on Earth, revealed the presence of glycine, alanine, and glutamic acid. Initial reports stated that racemic mixtures of these amino acids were found, suggesting no terrestrial contamination (*20*); however, further research revealed that glycine, alanine, glutamic acid, leucine, aspartic acid, and proline were present in an enantiomeric excess of near 33% (*21*). Regardless of the source of the non-biological molecules that transitioned into self-replicating life, evidence suggests that these molecules evolved to become capable of self-replication (*22*).

The technique used to detect chiral molecules in meteorites and analyze soil from the Viking Mars landing was gas chromatography (*15, 20, 23-25*). A disadvantage of using a GC column is its limitation to identifying only known biochemical compounds, as standards are necessary to determine retention times. Thus, in a given sample, if homochiral molecules, which are not known to be critical for Earth based

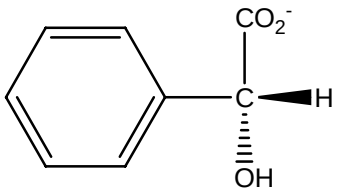
life, were present, they would likely be overlooked or mistaken for another compound as their chirality would not be detected by the instrument.

In 1996, a consortium of European and American scientists made the determination that homochirality is a characteristic sign of life and suggested that the search for extraterrestrial life should be renamed the Search for ExtraTerrestrial Homochirality (SETH) (*12*). Because optical rotation is superior to gas chromatography separation with regard to measuring possible extraterrestrial chiral molecules, so it would be of great benefit to develop a polarimeter for space missions (*12, 26*). This measure would eradicate the necessity to make a preconception about the extraterrestrial biochemistry (*12*).

If polarimetry indicates optical rotation in an extraterrestrial sample, this finding alone will conclusively show the presence of chiral molecules. Also, this instrument could identify a change in optical rotation over time, which could suggest one of two possibilities. Some organic compounds, with pK_a values in the range between 1 and about 12, will racemize with acid/base catalysis, through tautomerism, other organic compounds with pK_a values larger than 20 do not easily racemize, except through enzymatic catalysis. Therefore, finding a chiral compound involved in a racemization reaction would be indicative as a marker of the presence of life. Racemization, the conversion of an optically active substance to its racemic form, is a first-order reaction, thus the rate constant is

independent of concentration, hence even a small amount of substrate still results in a measurable rate. Also, racemization is most commonly associated with biotic systems: it helps organisms utilize both enantiomers of a molecule, which acts as a carbon or nitrogen source, thus aiding survival. Since only a biological molecule would likely be responsible for such a change, a reaction such as those carried out by racemase enzymes is another sign of the existence of life. This type of enzymatic reaction is explored in this thesis.

The objective of this thesis is to develop the enzyme mandelate racemase (MR) as a benchmark for exploration of extraterrestrial life. Understanding the behavior of this enzyme under conditions similar to those in extraterrestrial locations, including low temperatures and low water content, will elucidate the possibilities of proteins or similar molecules, which may exist in these locations. These studies can occur without an expensive extraterrestrial space mission. We have chosen as our test case the reaction catalyzed by the enzyme mandelate racemase, for reasons discussed in Chapter 2. In short, the substrate is a simple, inexpensive chiral molecule (Figure 1-1). The enzyme's crystal structure has been solved (*27-30*) and extensive studies on the mechanism (*27, 29, 31-33*), kinetics (*34-36*), and thermodynamics (*35, 36*) have been carried out in previous work. Most importantly, this enzyme performs a reaction that can be detected by a polarimetric assay.



(R)-Mandelate

Figure 1-1. Mandelate, substrate catalyzed by mandelate racemase.

Our first aim of the study, discussed in Chapter 3, involves the cloning of the mandelate racemase gene from *Pseudomonas putida*, with subsequent overexpression in *Escherichia coli*, purification, and characterization of MR using a polarimetric assay. This enzyme has been previously characterized using a circular dichroism assay as well as a coupled assay using mandelate racemase/mandelate dehydrogenase, and results of the present characterization studies are then compared to previously published values (*32, 34-38*). As part of the characterization, we also explored the irreversible thermal denaturation of mandelate racemase using differential scanning calorimetry (DSC), as discussed in Chapter 4. As the enzyme is octameric, we use DSC to determine the dissociation of the protein as well as the unfolding characteristics. The thermostability of many enzymes has been explored (*39-46*); however, this is the first time an enzyme of this size (octameric) has been studied by these methods.

Using the pure enzyme, we are then able to explore parameters that are important in the search for extraterrestrial life. This is the second aim of the thesis and is discussed in Chapter 5. All of the planets and moons that are explored for life feature temperatures lower than those on Earth, so it is important to test this reaction at low temperatures. Exploring the temperature limitations of the enzyme allows us to investigate conditions that lead to an inactive enzyme; hence, we would like to explore the activity of MR at temperatures below 0°C. Previous low temperature studies have been conducted for mechanistic studies of enzyme intermediates (*47, 48*). By slowing down the reaction one can determine enzymatic intermediates (*49, 50*). Enzymatic reactions at low temperatures have also been studied for the enhancement of food shelf life (*39, 51*).

Evaluation of enzymatic activity in non-aqueous media is the third aim of the study. Although water was or is most likely present on some planets and moons in small amounts (see Chapter 2), these environments chiefly consist of non-aqueous substances. In order to evaluate these conditions, we tested the MR enzymatic reaction in low-water environments. A variety of water-miscible as well as water-immiscible organic co-solvent systems are explored in Chapter 5. We specifically look at water-in-oil microemulsions, which have been implicated as possibly playing a role in the origin of life (*52*). We also employ

concentrated ammonium salt solutions which are known to be present in extraterrestrial locations (*53, 54*). Both of these systems have been explored with other enzymes, but neither has been used for mandelate racemase activity nor explored down to the lowest possible temperature.

The fourth aim of the project, which is discussed in Chapter 6, involves the investigation of the evolution of mandelate racemase. Demonstrating the ability of this enzyme to evolve can lead to an understanding of its ancestors. This enzyme is a member of the α/β barrel family, proteins that are quite unique in that the amino acid residues responsible for stability are separate from those responsible for catalysis. Specifically, mandelate racemase is a member of the enolase superfamily which has four subfamilies: muconate lactonizing enzyme (MLE), mandelate racemase (MR), 3-methylaspartate ammonia lyase (MAL) and enolase. Previously, enzymes included in this family have been easily evolved (i.e. with one or two amino acid changes) into enzymes with different functions (*55*). However, these changes all occurred within individual subfamilies. The goal of this part of the project is to evolve MR toward an N-acetylamino acid racemase (NAAR), both members of two different subfamilies. Such a goal has not been accomplished before. Finding an MR variant with simultaneous cross-reactivity for N-acylamino acid racemization gives rise to a possible ancestor of the enolase

superfamily. This, in turn, provides one more piece of information toward understanding the origin of life.

In summary, significant advances have recently been made in the investigation of the possibility of extraterrestrial life (*2, 9, 10, 12, 56-58*). Finding homochiral molecules would surely support the case for the existence of life (*12, 15, 59*), although searching for a racemate reaction would provide even better evidence. Therefore, the premise of this work is that the detection of a racemic reaction would be an excellent piece of evidence in the search for extraterrestrial life.

Using mandelate racemase as a test case, we will explore effects of low temperatures and thermostability, as well as non-aqueous media, such as those possibly found on Mars, Enceladus, Europa, and Titan, on enzymatic activity. Also, enhancing the substrate specificity of the enzyme toward the ancestor will provide insight into evolution of enzymes, allowing us to provide further evidence toward the understanding of the development of enzymes since their origination.

1.1 REFERENCES

(1) Wolfram, S. (2002) *A New Kind of Science*, Wolfram Media, Champaign, IL.

(2) Des Marais, D. J., Allamandola, L. J., Benner, S. A., Boss, A. P., Deamer, D., Falkowski, P. G., Farmer, J. D., Hedges, S. B., Jakosky, B. M., Knoll, A. H., Liskowsky, D., Meadows, V. S., Meyer, M. A., Pilcher, C. B., Nealson, K. H., Spormann, A. M., Trent, J. D., Turner, W. W., Woolf, N. J., and Yorke, H. W. (2003) The Nasa Astrobiology Roadmap. *Astrobiology 3*, 219-235.

(3) Cleland, C. (2006), Medical News Today.

(4) Cleland, C. E., and Chyba, C. F. (2002) Defining Life. *Origins of Life and Evolution of the Biosphere 32*, 387-393.

(5) Schrodinger, E. (1945) *What Is Life?*, Cambridge Univ. Press.

(6) Monod, J. (1970) *Chance and Necessity*, London.

(7) Feinberg, G., and Shapiro, R. (1980) *Life Beyond Earth*, William Morrow and Col., New York, NY.

(8) Shapiro, R., and Feinberg, G. (1995) in *Extraterrestrials: Where Are They?*, Cambridge U. Press, Cambridge.

(9) Koshland, D. E. (2002) The Seven Pillars of Life. *Science 295*, 2215-2216.

(10) Bada, J. (2004) How Life Began on Earth: A Status Report. *Earth and Planetary Science Letters 226*, 1-15.

(11) Bell, J. (2006) The Red Planet's Watery Past. *Scientific American*, 62-69.

(12) Macdermott, A. J., Barron, L. D., Brack, A., Buhse, T., Drake, A. F., Emery, R., Gottarelli, G., Greenberg, J. M., Haberle, R., Hegstrom, R. A., Hobbs, K., Kondepudi, D. K., Mckay, C., Moorbath, S., Raulin, F., Sandford, M., Schwartzman, D. W., Thiemann, W. H. P., Tranter, G. E., and Zarnecki, J. C. (1996) Homochirality as the Signature of Life: The Seth Cigar. *Planetary and Space Science 44*, 1441-1446.

(13) Watson, J. D., and Crick, F. H. (1953a) Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid. *Nature 171*, 964-967.

(14) Watson, J. D., and Crick, F. H. (1953b) Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature 171*, 737-738.

(15) Bada, J. L., Glavin, D. P., Mcdonald, G. D., and Becker, L. (1998) A Search for Endogenous Amino Acids in Martian Meteorite. *Science 279*, 362-365.

(16) Bernstein, M. P., Dworkin, J. P., Sandford, S. A., Cooper, G. W., and Allamandola, L. J. (2002) Racemic Amino Acids from the Ultraviolet Photolysis of Interstellar Ice Analogues. *Nature 416*, 401-403.

(17) Cohen, B. A., and Chyba, C. F. (2000) Racemization of Meteoritic Amino Acids. *Icarus 145*, 272-281.

(18) Munoz-Caro, G. M., Meierhenrich, U. J., Schutte, W. A., Barbier, B., Arcones-Segovia, A. A., Rosenbauer, H., Thiemann, W. H. P., Brack, A., and Greenberg, J. M. (2002) Amino Acids from Ultraviolet Irradiation of Interstellar Ice Analogues. *Nature 416*, 403-406.

(19) Skelley, A. M., Scherer, J. R., Aubrey, A. D., Grover, W. H., Ivester, R. H. C., Ehrenfreund, P., Grunthaner, F. J., Bada, J. L., and Mathies, R. A. (2005) Development and Evaluation of a Microdevice for Amino Acid Biomarker Detection and Analysis on Mars. *Proc. Natl. Acad. Sci. 102*, 1041-1046.

(20) Kvenvolden, K., Lawless, J., Pering, K., Peterson, E., Flores, J., Ponnamperuma, C., Kaplna, I. R., and Moore, C. (1970) Evidence

for Extraterrestrial Amino Acids and Hydrocarbons in Te Murchison Meteorite. *Nature 228*, 923-926.

(21) Engel, M. H., and Nagy, B. (1982) Distribution and Enantiomeric Composition of Amino Acids in the Murchison Meteorite. *Nature 296*, 837-840.

(22) Gilbreath, W. P. (1980) in *Proceedings of the 1980 NASA/ASEE Summer Study* (Robert A. Freitas, J., Ed.), NASA Conference Publication 2255, Santa Clara, CA.

(23) Powell, K. A., Ramer, S. W., Del Cardayre, S. B., Stemmer, W. P. C., Tobin, M. B., Longchamp, P. F., and Huisman, G. W. (2001) Directed Evolution and Biocatalysis. *Angew. Chem.-Int. Edit. 40*, 3948-3959.

(24) Rodier, C., Vandenabeele-Trambouze, O., Sternberg, R., Coscia, D., Coll, P., Szopa, C., Raulin, F., Vidal-Madjar, C., Cabane, M., Israel, G., Genier-Loustalot, M. F., Dobrijevic, M., and Despois, D. (2001) Detection of Martian Amino Acids by Chemical Derivation Coupled to Gas Chromatography: In Situ and Laboratory Analysis. *Space Life Sciences: Life in the Solar System: Prebiotic Chemistry, Chirality, and Space Biology 27*, 195-199.

(25) Szopa, C., Coscia, D., Rodier, C., Sternberg, R., and Raulin, F. (2001) Development and Analytical Aspects of Gas Chromatography for Space Exploration. *LC-GC 14*, 114-120.

(26) Macdermott, A. (1997) Distinguishing the Chiral Signature of Life in the Solar System and Beyond. *ProcSPIE 3111*, 272-279.

(27) Kenyon, G. L., Gerlt, J. A., Petsko, G. A., and Kozarich, J. W. (1995) Mandelate Racemase: Structure-Function Studies of a Pseudosymmetric Enzyme. *Acc. Chem. Res. 28*, 178-186.

(28) Hasson, M. S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G. L., Babbitt, P. C., Gerlt, J. A., Petsko, G. A., and Ringe, D. (1998) Evolution of an Enzyme Active Site: The Structure of a New Crystal Form of Muconate Lactonizing Enzyme Compared with Mandelate

Racemase and Enolase. *Proc. Natl. Acad. Sci. U. S. A. 95*, 10396-10401.

(29) Kallarakal, A. T., Mitra, B., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the K166r Mutant. *Biochemistry 34*, 2788-2797.

(30) Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate Racemase and Muconate Lactonizing Enzyme Are Mechanistically Distinct and Structurally Homologous. *Nature 347*, 692-694.

(31) Kenyon, G. L., and Hegeman, G. D. (1970) Mandelic Acid Racemase from *Pseudomonas Putida*. Evidence Favoring a Carbanion Intermediate in the Mechanism of Action. *Biochemistry 9*, 4036-4042.

(32) Bearne, S. L., and Wolfenden, R. (1997) Mandelate Racemase in Pieces: Effective Concentrations of Enzyme Functional Groups in the Transition State. *Biochemistry 36*, 1646-1656.

(33) Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent Evolution in the Enolase Superfamily: The Interplay of Mechanism and Specificity. *Arch. of Biochem. and Biophys. 433*, 59-70.

(34) Fee, J. A., Hegeman, G. D., and Kenyon, G. L. (1974) Mandelate Racemase from Pseudomonas Putida. Subunit Composition and Absolute Divalent Metal Ion Requirement. *Biochemistry 13*, 2528-2532.

(35) St. Maurice, M., and Bearne, S. L. (2002) Kinetics and Thermodynamics of Mandelate Racemase Catalysis. *Biochemistry 41*, 4048-4058.

(36) St. Maurice, M., and Bearne, S. L. (2000) Reaction Intermediate Analogues for Mandelate Racemase: Interaction between Asn 197

and the α-Hydroxyl of the Substrate Promotes Catalysis. *Biochemistry 39*, 13324-13335.

(37) Fee, J. A., Hegeman, G. D., and Kenyon, G. L. (1974) Mandelate Racemase from Pseudomonas Putida. Affinity Labeling of the Enzyme by D,L-α-Phenylglycidate in the Presence of Magnesium Ion. *Biochemistry 13*, 2533-2538.

(38) Felfer, U., Strauss, U. T., Kroutil, W., Fabian, W. M. F., and Faber, K. (2001) Substrate Spectrum of Mandelate Racemase Part 2. (Hetero)-Aryl-Substituted Mandelate Derivatives and Modulation of Activity. *Journal of Molecular Catalysis B: Enzymatic 15*, 213-222.

(39) Reat, V., Finney, J. L., Steet, A., Roberts, M. A., Smith, J., Dunn, R., Peterson, M., and Daniel, R. M. (2000) Cryosolvents Useful for Protein and Enzyme Studies Below -100°C. *J. Biochem. Biophys. Methods 42*, 97-103.

(40) Hernandez-Arana, A., Rojo-Dominguez, A., Altamirano, M. M., and Calcagno, M. L. (1993) Differential Scanning Calorimetry of the Irreversible Denaturation of *Escherchia Coli* Glucosamine-6-Phosphate Deaminase. *Biochemistry 32*, 3644-3648.

(41) Kurganov, B. I., Lyubarev, A. E., Sanchez-Ruiz, J. M., and Shnyrov, J. L. (1997) Analysis of Differential Scanning Calorimetry Data for Proteins: Criteria for Validity of One-Step Mechanism of Irreversible Protein Denaturation. *Biophys. Chem. 69*, 125-135.

(42) Lyubarev, A. E., Kurganov, B. I., Burlakova, A. A., and Orlov, V. N. (1998) Irreversible Thermal Denaturation of Uridine Phosphorylase from *Escherichia Coli* K-12. *Biophys. Chem. 70*, 247-257.

(43) Manly, S. P., Matthews, K. S., and Sturtevant, J. M. (1985) Thermal Denaturation of the Core Protein of *Lac* Repressor. *Biochemistry 24*, 3842-3846.

(44) Michnik, A., Drzazga, Z., Kluczewska, A., and Michalik, K. (2005) Differential Scanning Microcalorimetry Study of the Thermal Denaturation of Haemoglobin. *Biophys. Chem. 118*, 93-101.

(45) Quesada-Soriano, I., Garcia-Maroto, F., and Garcia-Fuentes, L. (2006) Kinetic Study on the Irreversible Thermal Denaturation of *Schistosome Japonicum* Glutathione S-Transferase. *Biochim. et Biophys. Acta 1764*, 979-984.

(46) Sanchez-Ruiz, J. M. (1992) Theoretical Analysis of Lumry-Eyring Models in Differential Scanning Calorimetry. *Biophys. J. 61*, 921-935.

(47) Douzou, P. (1977) *Cryobiochemistry: An Introduction*, Academic Press, New York, NY.

(48) Travers, F., and Barman, T. (1995) Cryoenzymology: How to Practice Kinetic and Structural Studies. *Biochimie 77*, 937-948.

(49) Bragger, J. M., Dunn, R. V., and Daniel, R. M. (2000) Enzyme Activity Down to -100°C. *Biochimica et Biophysica Acta 1480*, 278-282.

(50) Vajda, T. (1999) Cryo-Bioorganic Chemistry: Molecular Interactions at Low Temperature. *Cell. Mol. Life Sci. 56*, 398-414.

(51) Privalov, P. L. (1989) Thermodynamic Problems of Protein Structure. *Annu. Rev. Biophys. Chem. 18*, 47-69.

(52) Walde, P. (2006) Surfactant Assemblies and Their Various Possible Roled for the Origin(S) of Life. *Origins of Life and Evolution of the Biosphere 36*, 109-150.

(53) Lorenz, R. D., and Shandera, S. E. (2001) Physical Properties of Ammonia-Rich Ice: Application to Titan. *Geophysical Research Letters 28*, 215-218.

(54) Brown, R. H., Clark, R. N., Buratii, B. J., Cruikshank, D. P., Barnes, J. W., Mastrapa, R. M. E., Baurer, J., Newman, S. M., Momary, T., Baines, K.

H., Bellucci, G., Capaccioni, F., Cerroni, P., Combes, M., Coradini, A., Drossart, P., Formisano, V., Jaumann, R., Langevin, Y., Matson, D. L., Mccord, T. B., Nelson, R. M., Nicholson, P. D., Sicardy, B., and Sotin, C. (2006) Composition and Physical Properties on Enceladus Surface. *Science 311*, 1425-1428.

(55) Schmidt, D. M., Mundorff, E. C., Doijka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary Potential of (B/a)8-Barrels: Functional Promiscuity Produced by Single Substitutions in the Enolase Superfamily. *Biochemistry 42*, 8387-8393.

(56) Cockell, C. (2002) Astrobiology-a New Opportunity for Interdisciplinary Thinking. *Space Policy*, In Press.

(57) Maurel, M. C., and Decout, J. L. (1999) Origins of Life: Molecular Foundations and New Approaches. *Tetrahedron 485*, 3141-3182.

(58) Tarter, J. (2001) The Search for Extraterrestrial Intelligence (Seti). *Annu. Rev. Astrom. Astrophys. 39*, 511-548.

(59) Becker, L., Mcdonald, G. D., and Bada, J. L. (1997) Biomarkers for Analysis of Martian Samples. *Workshop on Early Mars (1997)*.

CHAPTER 2

BACKGROUND: THE SEARCH FOR EXTRATERRESTRIAL LIFE USING HOMOCHIRALITY AND RACEMIZATION

2.1 Origin of Life

Astrobiology has been defined as the search for life in the Universe. Recently, this definition has been amended to include the investigation of 'the origin and evolution of life in the universe'(*1*). For that reason, astrobiology addresses three fundamental questions: i). How does life originate and evolve? ii). Does life exist in the universe? iii). What is the future of life in the universe? One possibility is that extraterrestrial life originates and evolves in same way that life on Earth has evolved. Therefore, if we can understand the emergence and evolution of terrestrial life, we can gain a better understanding of what to search for on extraterrestrial entities. The Earth formed approximately 4.5 billion years ago, with the oldest sedimentary rocks containing biological material dating to around 3.8 billion years ago (*1-3*). A large knowledge gap surrounds the nearly 1 billion years between the formation of the Earth

and the origin/evolution of the first life forms. Currently, there are two main controversial theories regarding the origin of life. The first describes the emergence of life from primordial oceans, the pre-biotic soup theory (*4*). The other theory, the panspermia hypothesis, states that the initial seed for life on Earth arrived from an extraterrestrial source (*5*) (although this theory still requires an origin of life event to have occurred somewhere else.) Modern theorists reject the panspermia theory and focus on biological precursors such as amino acids and nucleotides that could be generated in the abiotic environment (*6*).

Oparin and Haldane developed one of the first hypotheses on the formation of a pre-biotic soup independently in the 1920s. They proposed that hydrogen cyanide and formaldehyde could have spontaneously formed complex molecules (*6-8*). In the 1950s, Miller and Urey performed their renowned experiments generating 17 of the 20 amino acids by applying electrical discharges simulating lightning to a gaseous mixture of simple compounds, including methane (CH_4), ammonia (NH_3), hydrogen (H_2), and water (H_2O), all presumed to be present on prebiotic Earth (*9-12*). In the absence of ammonia only α-hydroxy acids will form, so, the concentration of ammonia is critical for the synthesis of amino acids (*4*). Further products of abiotic synthesis of amino acids that are relevant to life, have been subsequently identified. These identifications include the Miller/Urey experiments as well as others using modern analytical

techniques such as ion-exchange chromatography/mass spectrometry (*9-11, 13, 14*).

The discovery of the structure of deoxyribonucleic acid (DNA) by Watson and Crick in 1953 shattered the origin of life scenario proposed by Oparin and Haldane (*15, 16*). However, DNA could not have originated on its own because it depends on proteins for duplication. Therefore, it has been theorized that ribonucleic acid (*17*) may have acted as an enzyme, promoting its own replication (*6*). The "RNA world" hypothesis developed in mid-1980, proposed that RNA was the first life-form on Earth which is supported by RNA's ability to store, transmit, and duplicate, also because RNA can reproduce on its own, it can execute the roles of both DNA and protein (*18, 19*). The discovery of ribozymes, RNA molecules with enzymatic activity, demonstrated that RNA has the ability to act as an enzyme (*20*). Efforts to develop a self-replicating ribozyme system in the laboratory are under way (*21*). For example, a polymerase ribozyme has been developed which is able to bind a primer/template duplex and elongate the 3'-terminus of the primer by 3',5'-phosphodiester bonds in a template-dependent fashion (*21-23*). Even with this advance, a laboratory synthesis of a self-replicating RNA system has yet to be demonstrated (*21*). So far, the addition of only 14 nucleotides has been achieved but a self replicating system requires a ribozyme on the order of 189 nucleotides (the size of a polymerase ribozyme) (*21, 24*).

In the search for extraterrestrial life, it is important to understand the components necessary for the existence of Earth based life. First and foremost, water is the essential basis for the biochemistry of all biological systems known on Earth. Organic polymers, such as amino acids, are also necessary to carry out replication as well as biological catalysis (*4*). Energy flow is also essential for the origin of life. The major sources considered include solar radiation (which is difficult to harness) and chemical redox energy. Chemical redox energy is derived from interactions between the effluents of a reduced mantle and a more oxidized atmosphere (*25*). Considering that these components are crucial for the existence of life, we must determine the habitats and micro-habitats most likely to contain such molecules.

2.2 Regions of Solar System to Explore

When searching for extraterrestrial life, scientists look for regions in the solar system with conditions that may have been or are capable of supporting some type of life. The two vital components for life are liquid water and organic polymers, such as peptides or proteins. Therefore, discovery of either of these would suggest the presence of life. In turn, it is then important to explore regions where the presence of these components is possible. The region of space where the conditions are favorable for the creation of life is referred to as the habitable zone (Figure 2-1). In this region temperatures are such that it would be possible

to find liquid water (*26*). According to data recently collected on various NASA missions, locations, which meet these criteria, include Mars, Europa, Enceladus, and Titan.

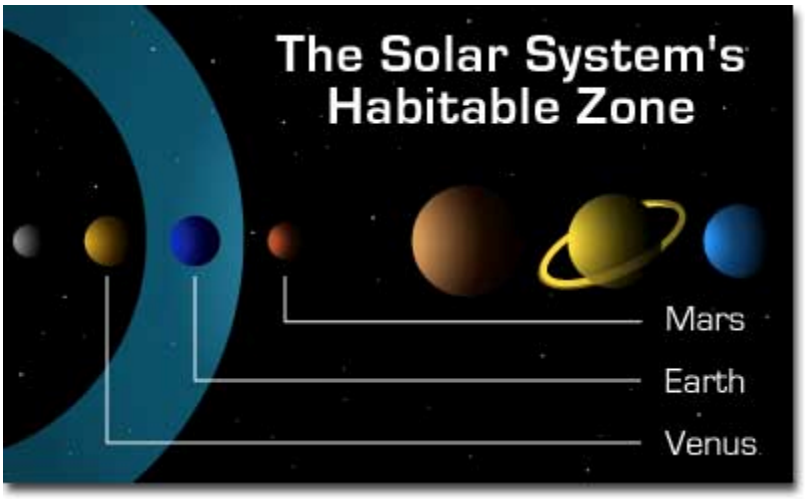


Figure 2-1. The habitable zone in the Solar System, represented by the blue band. Image from www.nasa.gov, accessed 03/07.

2.2.1 Mars

The surface of Mars is slightly colder than Earth with an equatorial temperature of 25 °C, to below -80°C at night (*27*) . Figure 2-2 gives a detailed look of the Eastern foot of "Burns cliff," which looks worn and suggests the flow of water at some point in the planet's history. Hematite (Fe_2O_3) has also been found, which often but not always forms in the presence of water (*28*). Recently, goethite (hydrated iron oxide-FeO(OH)), a mineral that can only form in the presence of water, has been discovered (*27, 29*). Further verification for the presence of water at

some point in the plane's history is with the detection of gypsum ($CaSO_4 \bullet 2\ H_2O$). Gypsum forms in the presence of water under highly evaporative conditions (*29*). The Mars exploration rovers, Spirit and Opportunity in combination with Mars Odyssey and Global Surveyer orbitors, have discovered the most concrete evidence for liquid water on Mars. Samples of soil obtained by the Mars rover Spirit sampled soil and found that its composition is more than 50% salt, evidence of the presence water because salt deposits as water evaporates (*27-30*). This evidence of salt allows for the hypothesis that biological molecules may be trapped in the salt crystals found in these soil samples (*27*). Also, high-resolution images identify regions featuring signs of liquid water runoff such as craters and erosion.

Figure 2-2. Detailed region of Mars, known as "Burns Cliff." Image http://www.nasa.gov/vision/universe/solarsystem/mer-121304b.html, accessed 04/05

2.2.2 Europa

Europa is a moon of Jupiter (Figure 2-3). The Hubble space telescope has identified a thin layer of atmospheric molecular oxygen and free water under the ice surface (*31, 32*). However, it appears that non-biological processes involving hydrogen peroxide produce this oxygen atmosphere. The exact mechanism by which it is produced is still unclear (*33*). Jupiter has an intense magnetic field that traps dust and charged particles. These charged particles then impact Europa's icy surface with enough force to produce water vapor (*34*). The gaseous water molecules then undergo reactions resulting in the formation of molecular hydrogen and oxygen (*36*).

Possible habitats on Europa include the icy layer, brine ocean, and seafloor environments (*37*). The average temperature of Europa is –153°C,

Figure 2-3. Surface of Europa. Image: www.cnn.com/TECH/ 9701/17/europa accessed 04/05

and model simulations suggest the presence of hypothetical oceans at temperatures around -20°C and an intensive radiation environment (*35, 36*). Due to its extremely cold temperature and intense radiation the icy surface layer is least likely to support life, but there is a much greater possibility below the ice surface. Simulations also suggest that a highly saline and acidic environment may exist under the ice (*37*).

2.2.3 Enceladus

The purpose of the joint NASA-ESA Cassini-Huygens mission is to gain a better understanding of the planet Saturn, along with its rings and moons (*38-40*). Recently, the Cassini spacecraft detected water/gas geysers on Enceladus (Figure 2-4), a moon of Saturn (*41*). The water geysers, which are located in the south polar region of the moon, are powered by gaseous carbon dioxide and due to the low gravity shoot thousands of kilometers into space (*42*).

The surface of Enceladus is just as interesting as its water geysers. There are deep canyons as well as variations in crater density, suggesting that there has been episodic resurfacing by volcanic action, a heat source (*42*). There are also mountains of amorphous and crystalline water ice, especially in the South Polar Regions (*43*). As the average surface temperature is about -201°C, the water geysers have to be powered from a heat source beneath the surface, such as volcanos (*43, 44*). The source

of liquid water in these geysers provides an excellent environment to explore in the search for extraterrestrial life.

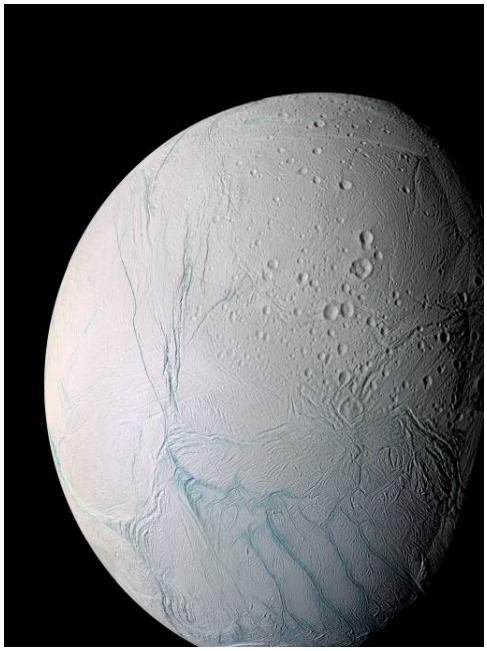


Figure 2-4. Saturn's moon Enceladus. Image: http://www.nasa.gov/mission_pages/cassini/main/index, accessed 04/05

2.2.4 Titan

Titan, a moon of Saturn, has a nitrogen atmosphere with a fog of hydrocarbons (Figure 2-5) (*45*). The chemical composition of Titan resembles early Earth but it is much colder, registering -178°C on the surface. Cassini scientists have identified radar-dark patches, and it has been hypothesized that these are liquid hydrocarbon lakes (*46*). Recently, the Huygens probe has revealed a terrain that resembles Earth with hills, liquid hydrocarbon river systems and mud flats. Also, methane is

abundant; it is constantly destroyed by sunlight and turned into chemical smog, so there must be a methane source (*45, 47, 48*). Water vapor has also been observed in the atmosphere (*49*).

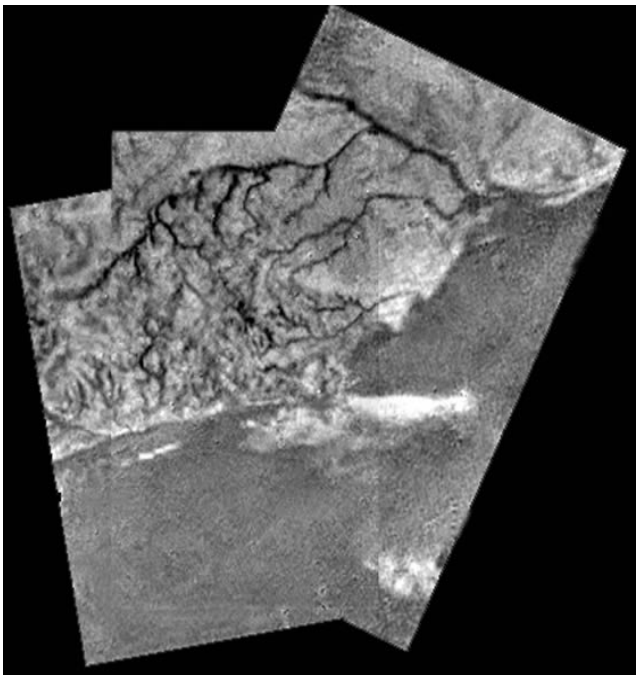


Figure 2-5. Hills and methane rivers on Titan. Image: ESA/NASA/JPL/University of Arizona, accessed 04/05

2.3 Origin of Homochirality

With the combination of searching for extraterrestrial life using homochiral molecules and searching for the origin of life, one may contemplate the origin of homochirality. Both biotic and abiotic theories regarding this phenomenon have been suggested, but as of yet no specific hypothesis has been agreed upon.

One approach is to first investigate if homochirality originated in one compound or molecular assembly before transferring to other molecules (*13, 50, 51*). It has been proposed that this process takes place through three, not necessarily sequential, events (*13*) such as mirror-symmetry breaking, enantiomeric enrichment, and chiral transmission (*52*). Mirror-symmetry breaking, considered to be the key step in the origin of homochirality, is defined as the process through which a small imbalance of the enantiomers by physical, chemical and/or statistical processes (*53*). Enantiomeric enrichment, the build-up of an enantiomeric excess of primitive molecules and has been shown to occur in polymerization reactions, in which the polymer preferentially forms homochiral units (*54*). Chiral transmission is the transfer of chirality from one molecule or molecular assembly to another (*13*). This process has been observed in clusters of serine molecules (*13, 52*).

The Vester-Ulbricht hypothesis, put forth by a group from Yale University in 1957, proposes that the generation of abiotic chiral organic products is possible through the creation of circularly polarized Bremsstrahlung, photons developed from longitudinally polarized electrons which are produced by parity violation during β-decay (*55, 56*). Many groups including Vester and Ulbricht (*56-61*) attempted to experimentally prove this hypothesis; however, these attempts have been unsuccessful. One set of experiments done at the Oak Ridge National

Laboratory involved 10.9 years of irradiation on several amino acid racemic mixtures, with ^{90}Sr-^{90}Y as the β-ray source (*58*). Analysis by gas chromatography did not provide evidence of enantiomeric excess (*58, 62*). Another set of experiments exposed DL-tryptophan to Bremsstrahlung from β-emitting ^{32}P, but again neither enantiomer was in excess (*60, 63*). Upon compilation of all the results, it has been determined that the Vester/Ulbricht hypothesis will not lead to the mechanism responsible for creating chiral molecules (*64*).

In 1966, Yamagata proposed that there must be a parity violating energy difference (PVED) between the D- and L-enantiomers (*65*). Using this idea, many investigators provided theoretical models which suggested that the energy difference between the enantiomers will be too small to measure experimentally (*66-69*). With that determination, it has been concluded that "there is no causal connection between parity violation in terrestrial biopolymers and nuclear processes" (*64*). In either case, the determinations made from the previous stated proposed hypotheses lead to the idea that the origin of chirality came from abiotic sources.

2.4 Racemization and the Search for Extraterrestrial Life

Although it is possible to search through extraterrestrial samples for chiral molecules, as in the Rosetta mission (*70*), the results are not conclusive as to the presence of extraterrestrial life. This is because

abiotic processes can also make this type of molecule. A better indication of life would be an excess of one enantiomer of a specific molecule. Observing this would require the measurement of a chiral molecules optical rotation however, yielding a static value, which is less useful than the dynamic measurement of a reaction rate, and would require the separation of the molecules in the sample. For life to exist there must be metabolic reactions occurring, and thus directly detecting the rate of such a reaction is our primary goal. Observing a metabolic reaction, like a racemization reaction, which results in a change in chirality over time would be particularly useful. Such a reaction would be less likely to occur abiotically, detecting it would not require separation of the sample, and most importantly its existence would strongly indicate the presence of a metabolic process and therefore that life is present.

2.5 The Model System

Exploring a reaction catalyzed by a racemase can be regarded as an indicator of the presence of life. The test case that we choose should be chemically and mechanistically straight-forward, well-studied, and have the possibility of a polarimetric assay. Also, because initial studies will be completed on a standard polarimeter the test reaction requires a large sample volume, hence the substrate should be easily obtained as well as inexpensive.

With these considerations, mandelate racemase (MR) makes an excellent test case for this system for many reasons. This enzyme catalyzes the racemization for the substrate mandelate (Figure 2-6). First, mandelate is one of the simplest chiral molecules. Also, mandelate has a very large specific optical rotation [α_D^{20}] = -152° (measured at 13.1 mM in water), which renders it ideal for use in a polarimetric assay.

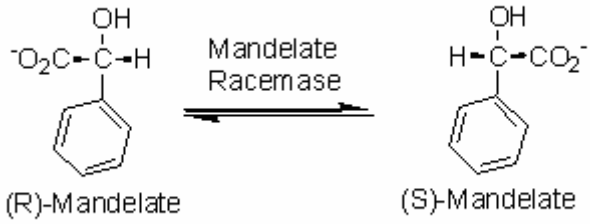


Figure 2-6. Reaction catalyzed by mandelate racemase (MR).

The enzymatic reaction using MR is fast (k_{cat} = 700 s^{-1}) while the uncatalyzed reaction is slow, the k_{uncat} value is $3.05*10^{-13}$ s^{-1}; the acceleration ratio k_{cat}/k_{uncat} equals $2*10^{15}$ s^{-1} (*71*). The uncatalyzed reaction is slow because the pK_a value of the α-hydrogen is 22 and 29 for mandelic acid and mandelate anion, respectively (*72*). With the uncatalyzed reaction 15 orders of magnitude slower than the catalyzed reaction (pH 7.5), observation of mandelate racemization very likely indicates biological catalysis and therefore supports the existence of life. After having determined a model system with many favorable

characteristics, it is important to explore more deeply the characteristics of the MR enzyme involved. The following discusses important features of the enzyme.

2.6 MR Background

2.6.1 α/β Barrel Enzymes

Mandelate racemase belongs to the structure class of α/β-barrel enzymes. The α/β-barrel domain has been found in approximately 10% of all enzymes with solved crystal structures (*73*). This highly similar structural group of enzymes catalyze many different types of reactions (*74*), including oxidation-reduction, hydrolysis, aldol cleavage, and cycloisomerization, to name a few (*75*).

2.6.1.1 The Structure of α/β-Barrel Enzymes

The α/β-barrel protein fold is composed of eight α-helices alternating the peptide backbone with eight β-strands, which are arranged in a barrel configuration. These alpha-helices then connect the parallel strands of the beta-sheets so that the N-terminus of each α-helix is adjacent to the C-terminal ends of the two neighboring β-strands (Figure 2-7), which are numbered sequentially from the N-terminus. The overall secondary structure is referred to as $(\beta\alpha)_8$. Most of the enzymes that contain the α/β-barrel structure feature an elliptical shape with the major axis of the ellipse pointing at various β-strands.

The active site of most α/β-barrel enzymes is found at the C-terminal end of the beta barrel with the active site residues residing in a pocket formed by the loops connecting the β-strand to the α-helix. The substrate

Figure 2-7. Model $(\alpha/\beta)_8$-barrel structure of the archetypical enzyme, TIM.

binding areas tend to be in the barrel. The βα-loops are those loops that follow the β-strands. It is believed that these are important for function, while the loops following the α-helix are important for stability (*76*). The βα-loops tend to be the most hydrophobic region of the barrel, and tends to be larger than the αβ-loops.

Most of the α/β-barrel enzymes feature additional structural elements in addition to the $(\alpha\beta)_8$ motif. These specific characteristics

separate the large group of α/β-barrel proteins into families. There are currently six families and multiple "cousins" with structures that are related to the α/β-barrel structure (Table 2-1). However, not all characteristics apply to each member of the family (*74*). These relatives tend to be very similar to the α/β-barrel motif, and differ by missing, distorted or added beta-strands or alpha-helices.

2.6.1.2 Evolution of α/β-Barrel Enzymes

The evolution of α/β-barrel enzymes has been a controversial topic. There are two possible pathways involved in the evolution of this family of proteins, convergent or divergent evolution, with evidence supporting both types. Convergent evolution states that the α/β-barrel structure evolved multiple times through enzymes of unrelated ancestry, that is nature converged on the α/β-barrel solution from multiple starting points (*74*). On the other hand, divergent evolution proposes that these enzymes share a common ancestry because they have a related structure (*74*).

One argument in support of convergent evolution suggests that, because the α/β-barrel motif is stable, enzymes tend to evolve into this fold. This is further supported by the fact that many enzymes in this group lack a common or even related chemical function. Another support of convergent evolution is the lack of sequence homology among the group

as a whole, where amino acid sequence identities tend to be in a range of 15-30% (*77*).

Table 2–1. Alpha/Beta Barrel Enzyme Families.

Family	Characteristics	Examples
A	Barrels nearly circular, Small helix between β-strand 8 and α-helix 8 Domain covering N-terminus of the barrel	Glycolate oxidase, Trimethylamine dehydrogenase
B	Major axis near β-strand 1 MR, MLE are missing final α-helix Domains cover the C-terminal end of the barrel Domain blocks N-terminus of the barrel	Mandelate racemase, Muconate cycloisomerase, Xylose isomerase, Mandelate Racemase
C	Composed of a single domain Major axis near β-strand 3 Small helix between β-strand 8 and α-helix 8	Tryptophan synthase, Triose phosphate isomerase (TIM)
D	Have an α-helix preceding the alpha/beta barrel This helix blocks the N-terminal end of the barrel	Fructose biphosphate aldolase, 2-keto-3-deoxy-6-phosphogluconate
E	Major axis of ellipse point toward β-strand Large loops at the carboxyl-terminal end of the barrel after β-strand 4 and 7 Bind NAD	Aldose reductase, 3α-hydroxysteroid dehydrogenase
F	Long thin barrel structure Group of alpha helices in the carboxyl-terminal loop between β-strand 1 and α-helix 1	Phosphotriesterase, Adenosine deaminase

However, there is higher sequence homology among individual families and it may suggest alternately that either the evolution is indeed convergent and the different families represent the different lineages that have converged on the similar scaffold, or simply that the ancestral enzyme is very old and thus enzymes diverged into different families long enough ago to significantly reduce their homology (*74, 78*).

There is further evidence to support divergent evolution: the motif is accounts for 10% of all known enzymes (*73, 79*), there is similar location of the active site (*74*), also these enzymes from this group can be separated into families using structural similarities, function, and sequence homology (*74*). Divergent evolution could also be supported if gene duplication from a pre-existing enzyme is necessary for catalytic diversity to evolve. By studying the evolution of the mandelate pathway, Petsko et. al. determined that gene duplication is possible (*80*). Observing the tendency for metabolic pathways to evolve in a reverse fashion, this group studied the relationship between each enzyme in the mandelate pathway. Through crystal structure analysis of MR the group discovered a very strong structural similarity to MLE, an enzyme in the β-ketoadipate pathway downstream of the mandelate pathway (*81*). The structure of MLE also from *P. putida* (*82*) superimposes to within less than 1.3 Å on the structure of MR (*81*). The similarities between these structures suggest that the gene that encodes for MLE was duplicated/and mutated to form MR.

O'Brien and Herschlag proposed that α/β-barrel enzymes evolved through catalytic promiscuity (*83*). Using the idea that enzymes evolve through gene duplication of a pre-existing enzyme, they suggest that α/β-barrel enzymes have the ability to catalyze additional reactions at a lower rate. This would explain why α/β-barrel enzymes have such similar structure but catalyze such a wide variety of reactions. It has been proposed that the α/β-barrel enzyme's function evolves faster than its sequence as a whole which in turn evolves faster than the overall structure (*78*), further support for divergent evolution. As has been shown, there is evidence supporting both possible types of evolution of α/β-barrel enzymes. However, the more convincing evidence supports divergent evolution.

2.6.2 Enolase Superfamily

Mandelate racemase is a member of the enolase superfamily (*84*). The enolase superfamily belongs to Family B of the α/β barrel group of proteins and consists of four subfamilies: mandelate racemase (MR), muconate lactonizing enzyme (MLE), 3-methylaspartate ammonia lyase (MAL), and enolase (*77*). Enzymes of the enolase superfamily contain two domains, a mixed α/β domain formed by the N- and C-termini (capping domain), and a larger $(\beta/\alpha)_7\beta$ barrel domain. Many differences between family members are due to variability in the sequences and structures of the capping domain.

All members of this family have a binding site for a divalent metal ion formed by three-amino acids, located at the ends of the third, fourth, and fifth β strands, which are D195, E221, and E247 in MR. This metal ion, Mg^{2+} in MR, then assists in the abstraction of the α-proton of a carboxylic acid to form an enolic intermediate. The family is ordered into subfamilies based on which amino acid residues are used to abstract or donate a proton to the carboxylate anion substrate (Table 2-2). All active site residues are found at the C-terminal ends of the beta-strand of the barrel domain.

Table 2-2. The subfamilies of the Enolase superfamily. (N) represents the number of the β-strand residue.

Subfamily	No. of Acid/Base residues	Example
Mandelate racemase (MR)	2	MR-Lys166(2^{nd})and H-bonded Asp270(6^{th})-His297(7^{th}) dyad
Muconate lactonizing enzyme (MLE)	2	MLE-Lys169(2^{nd})and Lys273(6^{th})
3-methylaspartate ammonia lyase (MAL)	1	3-methylaspartate ammonia lyase-Lys323 (6^{th})
Enolase subfamilies	1	Yeast Enolase-Lys345(6^{th})

The mandelate racemase (MR) subfamily members all have a conserved histidine residue that acts as a catalytic base. Mandelate racemase catalyzes the interconversion of (*R*)- and (*S*)-enantiomers of

mandelate, by using two acid/base catalysts to abstract the α-proton of the substrate. The metal binding residues for MR are D195, E221, and E247, while the catalytic residues are K166 as part of the KXK motif common to this superfamily, and the catalytic dyad D270/H297 (Figure 2-8).

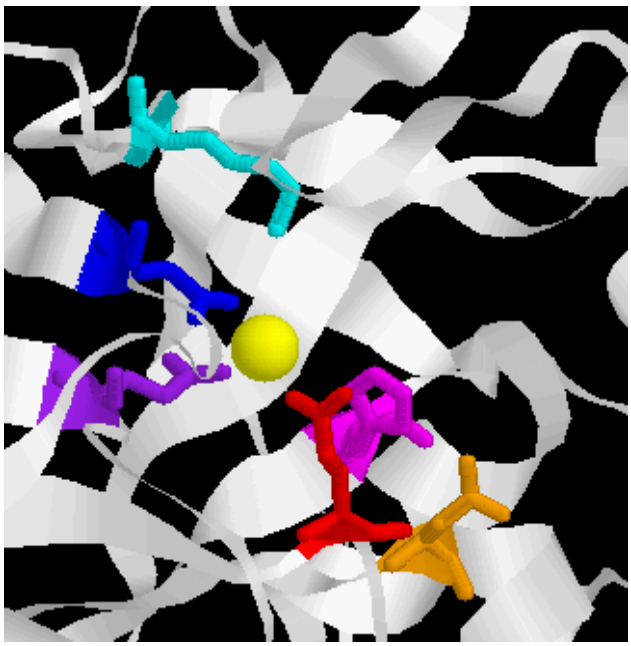


Figure 2-8. Active site of MR. Metal binding site: blue-D195, purple-E221, and red-E247. Active site: cyan-K166, magenta H297, orange-D270. Yellow-Mg^{2+}

2.7 Summary

Investigating the origins of life extra-terrestrially has become a topic of interest. Possible prebiotic chemicals have been detected on Mars, Europa, and Titan. Finding a homochiral molecule would surely be a basis for the generation of life, however, finding a change in chirality provides

more conclusive evidence for life. Therefore, this work involves the investigation of a reaction catalyzed by a racemic enzyme that is a superior piece of evidence for the sign of life.

Mandelate racemase has been suggested as the test case. Not only is mandelate a simple chiral molecule, but also it has a large specific optical rotation, and is thus easily detected in a polarimeter. The background reaction for MR is slow while the enzymatic reaction is fifteen orders of magnitude faster. There has been a significant amount of research focused on this enzyme, so a great deal is known.

Using MR as a test case, we will be able to explore conditions such as low temperatures and non-aqueous medium similar to those possibly found on Mars, Europa, and Titan. We would also like to investigate the evolution of the enolase family of enzymes. By broadening the substrate specificity toward another subfamily member, we can explore the possibilities for the ancestral enzyme for this superfamily, giving insight into the evolution of enzymes in general.

2.8 REFERENCES

(1) Cockell, C. (2002) Astrobiology-a New Opportunity for Interdisciplinary Thinking. *Space Policy*, In Press.

(2) Maurel, M. C., and Decout, J. L. (1999) Origins of Life: Molecular Foundations and New Approaches. *Tetrahedron 485*, 3141-3182.

(3) Mojzsis, S. J., Arrhenius, G., Mckeegan, K. D., Harrison, T. M., Nutman, A. P., and Friend, C. P. L. (1996) Evidence for Life on Earth before 3,800 Million Years Ago. *Nature 384*, 55-59.

(4) Bada, J. (2004) How Life Began on Earth: A Status Report. *Earth and Planetary Science Letters 226*, 1-15.

(5) Tarter, J. (2001) The Search for Extraterrestrial Intelligence (Seti). *Annu. Rev. Astrom. Astrophys. 39*, 511-548.

(6) Hiscox, J. A. (2001) An Overview of the Origin of Life: The Case for Biological Prospecting on Mars. *Earth, Moon and Planets 87*, 191-212.

(7) Oparin, A. I. (1938) *The Origin of Life*, Macmillan, New York.

(8) Haldane, J. B. S. (1928) *Possible Worlds, and Other Papers*, Harper & Brothers, New York.

(9) Miller, S. L. (1953) A Production of Amino Acids under Possible Primitive Earth Conditions. *Science 117*, 528-529.

(10) Miller, S. L. (1957) The Mechanism of Synthesis of Amino Acids by Electric Discharges. *Biochimica et Biophysica Acta 23*, 480-489.

(11) Miller, S. L., and Urey, H. C. (1959) Organic Compound Synthesis on the Primitive Earth. *Science 130*, 245-251.

(12) Urey, H. C. (1952) On the Origin of Continents and Mountains. *Science 38*, 351-363.

(13) Nanita, S. C., and Cooks, R. G. (2006) Serine Octamers: Cluster Formation, Reactions, and Implications for Biomolecule Homochirality. *Angewandte Chemie 45*, 554-569.

(14) Ring, D., Wolman, Y., Friedmann, N., and Miller, S. L. (1972) Prebiotic Synthesis of Hydrophobic and Protein Amino Acids. *Proc. Natl. Acad. Sci. 69*, 765-768.

(15) Watson, J. D., and Crick, F. H. (1953b) Genetical Implications of the Structure of Deoxyribonucleic Acid. *Nature 171*, 737-738.

(16) Watson, J. D., and Crick, F. H. (1953a) Molecular Structure of Nucleic Acids; a Structure for Deoxyribose Nucleic Acid. *Nature 171*, 964-967.

(17) Hernandez-Arana, A., Rojo-Dominguez, A., Altamirano, M. M., and Calcagno, M. L. (1993) Differential Scanning Calorimetry of the Irreversible Denaturation of *Escherchia Coli* Glucosamine-6-Phosphate Deaminase. *Biochemistry 32*, 3644-3648.

(18) Gilbert, W. (1986) The Rna World. *Nature 319*, 618.

(19) Woese, C. R. (2001) Translation: In Retrospect and Prospect. *RNA 7*, 1055-1067.

(20) Cech, T. R. (1993) The Efficiency and Versatility of Catalytic Rna: Implications for an Rna World. *Gene 135*, 33-36.

(21) Muller, U. F. (2006) Re-Creating and Rna World. *Cell. Mol. Life Sci. 63*, 1278-1293.

(22) Johnston, W. K., Unrau, P. J., Lawrence, M. S., Glasner, M. E., and Bartel, D. P. (2001) Rna-Catalysed Rna Polymerization: Accurate

and General Rna-Templated Primer Extension. *Science 292*, 1319-1325.

(23) Bartel, D. P., and Szostak, J. W. (1993) Isolation of New Ribozymes from a Large Pool of Random Sequences. *Science 261*.

(24) Curtis, E. A., and Bartel, D. P. (2005) New Catalytic Structures from an Existing Ribozyme. *Nat. Struct. Biol. 12*, 994-1000.

(25) Shapiro, R. (2006) Small Molecule Interactions Were Central to the Origin of Life. *The Quarterly Review of Biology 81*, 105-125.

(26) Raymond, S. N., Barnes, R., and Kaib, N. A. (2006) Predicting Planets in Known Extrasolar Planetary Systems. *Astrophysical Journal 644*, 1223-1231.

(27) Haskin, L. A., Wang, A., Jolliff, B. L., Mcsween, H. Y., Clark, B. C., Des Marais, D. J., Mclennan, S. M., Tosca, N. J., Hurowitz, J. A., Farmer, J. D., Yen, A., Squyres, S. W., Arvidson, R. E., Klingelhofer, G., Schroder, C., De Souza, P. A., Ming, D. W., Gellert, R., Zipfel, J., Bruckner, J., Bell, J. F., Herkenhoff, K., Christensen, P. R., Ruff, S., Blaney, D., Gorevan, S., Cabrol, N. A., Crumpler, L., Grant, J., and Soderblom, L. (2005) Water Alteration of Rocks and Soils on Mars at the Spirit Rover Site in Gusev Crater. *Nature 436*, 66-69.

(28) Wang, A., Haskin, L. A., Squyres, S. W., Jolliff, B. L., Crumpler, L., Gellert, R., Schroder, C., Herkenhoff, K., Hurowitz, J., Tosca, N. J., Farrand, W. H., Anderson, R., and Knudson, A. T. (2006) Sulfate Deposition in Subsurface Regolith in Gusev Crater, Mars. *Journal of Geophysical Research-Planets 111*.

(29) Clark, B. C., Morris, R. V., Mclennan, S. M., Gellert, R., Jolliff, B., Knoll, A. H., Squyres, S. W., Lowenstein, T. K., Ming, D. W., Tosca, N. J., Yen, A., Christensen, P. R., Gorevan, S., Bruckner, J., Calvin, W., Dreibus, G., Farrand, W., Klingelhoefer, G., Waenke, H., Zipfel, J., Bell, J. F., Grotzinger, J., Mcsween, H. Y., and Rieder, R. (2005) Chemistry and Mineralogy of Outcrops at Meridiani Planum. *Earth and Planetary Science Letters 240*, 73-94.

(30) Thomas, M., Clarke, J. D. A., and Pain, C. F. (2005) Weathering, Erosion and Landscape Processes on Mars Identified from Recent Rover Imagery, and Possible Earth Analogues. *Australian Journal of Earth Sciences 52*, 365-378.

(31) Nimmo, F., Prockter, L., and Schenk, P. (2005) Europa's Icy Shell: Past and Present State, and Future Exploration. *Icarus 177*, 293-296.

(32) Nimmo, F., and Giese, B. (2005) Thermal and Topographic Tests of Europa Chaos Formation Models from Galileo E15 Observations. *Icarus 177*, 327-340.

(33) Hudson, R. L., and Moore, M. H. (2006) Infrared Spectra and Radiation Stability of H2o2 Ices Relevant to Europa. *Astrobiology 6*, 483-489.

(34) Teolis, B. D., Loeffler, M. J., Raut, U., Fama, M., and Baragiola, R. A. (2006) Ozone Synthesis on the Icy Satellites. *Astrophysical Journal 644*, L141-L144.

(35) Mateo-Marti, E., Prieto-Ballesteros, O., Sobrado, J. M., Gomez-Elvira, J., and Martin-Gago, J. A. (2006) A Chamber for Studying Planetary Environments and Its Applications to Astrobiology. *Measurement Science & Technology 17*, 2274-2280.

(36) Ivanov, B. A. (2006) Earth/Moon Impact Rate Comparison: Searching Constraints for Lunar Secondary/Primary Cratering Proportion. *Icarus 183*, 504-507.

(37) Marion, G.M., Eicken, H., and Payne, M.C. (2003) The Search for Life on Europa: Limiting Environmental Factors, Potential Habitats, and Earth Analogues. *Astrobiology 3*, 785-811.

(38) (2006) Geysers on Saturn's Moon Enceladus. *American Scientist 94*, 219-219.

(39) (2006) The Possiblity for Life on Enceladus. *Astrobiology 6*, 239-239.

(40) Kargel, J. S. (2006) Perspective - Enceladus: Cosmic Gymnast, Volatile Miniworld. *Science 311*, 1389-1391.

(41) Spaun, F., Schmidt, J., Albers, N., Horning, M., Makuch, M., Seib, M., Kempf, S., Srama, R., Dikarev, V., Helfert, S., Moragas-Klostermeyer, G., Krivov, A. V., Sremcevic, M., Tuzzolino, A. J., Economou, T., and Grun, E. (2006) Cassini Dust Measurements at Enceladus and Implications for the Origin of the E Ring. *Science 311*, 1416-1418.

(42) Kargal, J. S. (2006) Enceladus: Cosmic Gymnast, Volatile Miniworld. *Science 311*, 1389-1391.

(43) Spitale, J. N., Jacobson, R. A., Porco, C. C., and Owen, W. M. (2006) The Orbits of Saturn's Small Satellites Derived from Combined Historic and Cassini Imaging Observations. *Astronomical Journal 132*, 692-710.

(44) Ostro, S. J., West, R. D., Janssen, M. A., Lorenz, R. D., Zebker, H. A., Black, G. J., Lunine, J. I., Wye, L. C., Lopes, R. M., Wall, S. D., Elachi, C., Roth, L., Hensley, S., Kelleher, K., Hamilton, G. A., Gim, Y., Anderson, Y. Z., Boehmer, R. A., and Johnson, W. T. K. (2006) Cassini Radar Observations of Enceladus, Tethys, Dione, Rhea, Iapetus, Hyperion, and Phoebe. *Icarus 183*, 479-490.

(45) Vuitton, V., Yelle, R. V., and Anicich, V. G. (2006) The Nitrogen Chemistry of Titan's Upper Atmosphere Revealed. *Astrophysical Journal 647*, L175-L178.

(46) Friedman, R., NASA Astrobiology Institute.

(47) Mitchell, R. T. (2006) Cassini/Huygens at Saturn and Titan. *Acta Astronautica 59*, 335-343.

(48) Karatekin, O., and Van Hoolst, T. (2006) The Effect of a Dense Atmosphere on the Tidally Induced Potential of Titan. *Icarus 183*, 230-232.

(49) Coustenis, A., Salama, A., Lellouch, E., Encrenaz, T., Bjoraker, G. L., Samuelson, R. E., Graauw, T. D., Feuchtgruber, H., and Kessler, M. F. (1998) Evidence for Water Vapor in Titan Atmosphere from Iso/Sws Data. *Astronomical Astrophysics 336*, L85-L89.

(50) Cooks, R. G., Zhang, D. X., Koch, K. J., Gozzo, F. C., and Eberlin, M. N. (2001) Chiroselective Self-Directed Octamerization of Serine: Implications for Homochirogenesis. *Analytical Chemistry 73*, 3646-3655.

(51) Popa, R. (1997) A Sequential Scenario for the Origin of Biological Chirality. *Journal of Molecular Evolution 44*, 121-127.

(52) Koch, K. J., Gozzo, F. C., Nanita, S. C., Takats, Z., Eberlin, M. N., and Cooks, R. G. (2002) *Angewandte Chemie 114*, 1797-1800.

(53) Macdermott, A. (1995) *Origins of Life Evol. Biosphere 25*, 191-199.

(54) Green, M. M., Cheon, K. S., Yang, S. Y., Park, J. W., Swansburg, S., and Liu, W. (2001) *Acc. Chem. Res. 34*.

(55) Vester, F., Ulbricht, T. L. V., and H., K. (1959) Optical Activity and Parity Violation in B-Decay. *Naturwiss. 46*, 68.

(56) Ulbricht, T. L. V., and Vester, F. (1962) Attempts to Induce Optical Activity with Polarized B-Radiation. *Tetrahedron 18*, 629-637.

(57) Garay, A. S. (1968) Origin and Rold of Optical Isomery in Life. *Nature 219*, 338-340.

(58) Bonner, W. A. (1974) Experiments on the Origin of Molecular Chirality by Parity Non-Conservation During B-Decay. *J. Mol. Evol. 4*, 23-29.

(59) Bernstein, W. J., Lemmon, R. M., and Calvin, M. (1972) An Investigation of the Possible Differential Radiolysis of Amino Acid Optical Isomers by 14c Betas, in *Molecular Evolution, Prebiological*

and Biological (Rohlfing, D. L. a. O., A.I., Ed.) pp 151-155, Plenum, New York.

(60) Darge, W., Laczko, I., and Thiemann, W. (1976) Stereoselectivity of B Irradiation of D,L-Tryptophan in Aqueous Solution. *Nature 261*, 522-524.

(61) Keszthelyi, L., and Vincze, I. (1975) Absorption of Circularly Polarized G-Radiation in L- and D-Amino Acids. *Radiat Environ Biophys 12*, 181-185.

(62) Bonner, W. A., and Liang, Y. (1984) B-Decay, Bremsstrahlen and the Origin of Molecular Chirality. *J. Mol. Evol. 21*, 84-89.

(63) Bonner, W. A. (1996) The Quest for Chirality, in *Physical Origin of Homochirality in Life* (Cline, D. B., Ed.) pp 17-49, Americna Institute of Physics, Woodbury, NY.

(64) Bonner, W. A. (2000) Parity Violation and the Evolution of Biomolecular Homochirality. *Chirality 12*, 114-126.

(65) Yamagata, Y. (1966) A Hypothesis for the Asymmetric Appearance of Biomolecules on Earth. *J. Theor. Biol. 11*, 495-498.

(66) Rein, D. W. (1974) Some Remarks on Parity Violating Effects of Intramolecular Interactions. *J. Mol. Evol. 4*, 15-22.

(67) Morozov, L. L., Kuz'min, V. V., and Goldanskii, V. I. (1983) Comparative Analysis of the Role of Statistical Fluctuations and Factor of Advantage (Parity Non-Conservation) in the Origins of Optical Activity. *Origins of Life 13*, 119-138.

(68) Mason, S. F., and Tranter, G. E. (1983) The Parity-Violating Energy Difference between Enantiomeric Molecules. *Chem. Phys. Lett. 94*, 34-37.

(69) Garay, A. S., and Ahlgren-Beckendorf, J. A. (1990) Differential Interaction of Chiral B-Particles with Enantiomers. *Nature 346*, 451-453.

(70) Macdermott, A. J., Barron, L. D., Brack, A., Buhse, T., Drake, A. F., Emery, R., Gottarelli, G., Greenberg, J. M., Haberle, R., Hegstrom, R. A., Hobbs, K., Kondepudi, D. K., Mckay, C., Moorbath, S., Raulin, F., Sandford, M., Schwartzman, D. W., Thiemann, W. H. P., Tranter, G. E., and Zarnecki, J. C. (1996) Homochirality as the Signature of Life: The Seth Cigar. *Planetary and Space Science 44*, 1441-1446.

(71) Wolfenden R, S. M. (2001) The Depth of Chemical Time and the Power of Enzymes as Catalysts. *Acc. Chem. Res. 34*, 938-945.

(72) Gerlt, J. A., Kozarich, J. W., Kenyon, G. L., and Gassman, P. G. (1991) *J. Am. Chem. Soc. 113*, 9667-9669.

(73) Gerlt, J. (2000) New Wine from Old Barrels. *Nat. Struct. Biol. 7*, 171-173.

(74) Farber, G. K., and Petsko, G. A. (1990) The Evolution of a/B Barrel Enzymes. *Trends Biochem Sci 15*, 228-234.

(75) Copley, R. R., and Bork, P. (2000) Homology among (Beta Alpha)(8) Barrels: Implications for the Evolution of Metabolic Pathways. *Journal of Molecular Biology 303*, 627-640.

(76) Wierenga, R. K. (2001) The Tim-Barrel Fold: A Versatile Framework for Efficient Enzymes. *FEBS Lett. 492*, 193-198.

(77) Reardon, D. A. G. K. F. (1995) The Structure and Evolution of α/B Barrel Proteins. *FASEB J. 9*, 297-503.

(78) Reardon, D., and Farber, G. K. (1995) The Structure and Evolution of α/B Barrel Proteins. *FASEB J. 9*, 297-503.

(79) Nagano, N., Hutchinson, E. G., and Thornton, J. M. (1999) Barrel Structures in Proteins: Automatic Identification and Classification Including a Sequence Analysis of Tim Barrels. *Protein Science 8*, 2072-2084.

(80) Petsko, G. A., Kenyon, G. L., Gerlt, J. A., Ringe, D., and Kozarich, J. W. (1993) On the Origin of Enzymatic Species. *Trends Biochem.Sci. 18*, 372-376.

(81) Neidhart Dj, K. G., Gerlt Ja, Petsko Ga. (1990) Mandelate Racemase and Muconate Lactonizing Enzyme and Mechanistically Distinct and Structurally Homologous. *Nature 347*, 692-694.

(82) Goldman, A., Ollis, D. L., and Steitz, T. A. (1987) Crystal Structure of Muconate Lactonizing Enzyme at 3 a Resolution. *J. Mol. Biol. 194*, 143-153.

(83) O'brien, P. J., and Herschlag, D. (1999) Catalytic Promiscuity and the Evolution of New Enzymatic Activities. *Chem. Biol. 6*, R91-R105.

(84) Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) The Enolase Superfamily: A General Strategy for Enzyme-Catalyzed Abstraction of the Alpha-Protons of Carboxylic Acids. *Biochemistry 35*, 16489-16501.

CHAPTER 3

AIM 1: CLONING, OVEREXPRESSION, PURIFICATION, AND CHARACTERIZATION MANDELATE RACEMASE FROM *Pseudomonas putida*

Racemization of mandelate catalyzed by mandelate racemase (MR) was chosen as the test case for the investigation of developing a benchmark in the search for extraterrestrial life. This enzyme has been extensively studied. The x-ray crystal structure (resolved to 1.8 Å) has been solved (*1*), and the catalytic mechanism determined (*2, 3*). Also, published kinetic and thermodynamic data are available (*4*). This chapter reports our results obtained from the cloning the gene for mandelate racemase from *P. putida* with subsequent overexpression of the protein in *E. coli,* and then purification and characterization of the MR protein.

3.1 Background

3.1.1 MR Protein Production

Initial studies on this enzyme began in the early 1970s with protein derived from *Pseudomonas putida*. Using primarily ammonium sulfate

precipitation and gel filtration the enzyme was purified 550-fold in six steps (*5*). MR protein is produced in *Pseudomonas putida* by inducing with both glucose and mandelate as carbon sources. The organism first preferentially utilizes the glucose until it is exhausted at which time the metabolism switches to mandelate indirectly causing the media to turn a grey/blue color change. This color is temporary, and is caused by a Fe(III)-complex with catechol which is produced during mandelate metabolism. The optimum time to harvest the cells was determined to be one hour after the grey/blue media returns yellow (*5*). This color shift occurs after the mandelate is exhausted and the catechol is metabolized.

Although pure protein can be produced using this method, using a recombinant gene allows for a higher yield because a stronger induction system can be included. Such a construct was developed in the late 1980s (*6*). The gene was cloned using a metabolic selection and its sequence analyzed. The recombinant protein was purified using a similar protocol as before, and when both proteins were compared they had similar specific activities. In this work, we purified the protein in both ways described above, as well as adding a histine tag to the gene, which has not been previously reported.

3.1.2 MR Assays

In the early 1970s, MR activity was measured using a spectrophotometric assay and coupled with the enzyme (*S*)-mandelate

dehydrogenase, which is the enzyme following MR in the mandelic acid pathway in *P. putida*. This enzyme catalyzes the oxidation of (*S*)-mandelate to benzoylformate while at the same time reducing 2,6-dichlorophenolindophenol which has a chromophore that absorbs in the oxidized form spectrophotometrically at 600 nm (*5*). However, some problems were encountered with this assay, most importantly the high turbidity of the mixture caused by the dehydrogenase (*7*).

To overcome this problem, in the late 1970s, Kenyon and Hegeman (1979) developed an assay for MR using circular dichroic (CD) measuments (*8*). CD is a direct assay that monitors changes in the substrate at 262 nm. Both mandelate enantiomers have molar ellipticities with the same magnitude, but with opposite signs, allowing changes to be used to follow the interconversion of the substrate. Another assay that allows direct measurement uses polarimetry, measuring the change in optical rotation over time (*9*).

Both CD and polarimetry are an improvement on the coupled assay because they are measuring racemization directly, however, neither of these methods are sensitive enough to accurately measure velocities at concentrations near or below the estimated K_M value (approximately 0.23 mM at pH 7.5 for R-mandelate) (*8*). St. Maurice and Bearne then developed a more sensitive assay (*7*). This high-performance liquid chromatography (HPLC) assay is a fixed time assay which quantifies

mandelate racemase activity by separation and detection of (*R*)- and (*S*)-mandelate concentrations with reverse-phase high-performance liquid chromatography using a chiral column and absorbance detection. Although this assay may have superior resolution, the equipment is expensive, the assay is time consuming and using an endpoint assay causes error when measuring rates. In this work we use the polarimetric assay.

3.1.3 Kinetics and Thermodynamics of MR

Since reaction conditions on the extraterrestrial planets and moon will be much different than on Earth it is important to understand the relationship between the kinetic and thermodynamic behavior of the enzyme under these extreme conditions, i.e. low temperatures and non-aqueous media. The activation energy of MR has been studied extensively (*10*).

The Kenyon group took initial kinetic measurements of the enzyme from *P. putida* and found that the value of the equilibrium constant, K_{eq}, was approximately equal to one, as is expected for a racemase reaction (*11*). The Kenyon group also measured kinetic data for the recombinant gene, this can be found in Table 3-1. The K_M value was found to be lower for the enzyme purified from *P. putida* when compared to that from the recombinant, while the k_{cat} values for each are similar. Both groups used the CD assay for measurement (*11*, *12*).

Table 3-1. Kinetic data for MR.

Substrate	K_m (mM)	k_{cat} (s^{-1})	Reference
(R)-mandelate	0.70 ± 0.06	526 ± 17	Recombinant (*11*)
(S)-mandelate	0.54 ± 0.06	467 ± 24	Recombinant (*11*)

In 1997, the primary kinetic isotope effect of MR-catalyzed deuterium incorporation into the α-position of mandelate was measured and found to be equal to about 3.2 (*10*). The first order rate constants for racemization of mandelate in H_2O and in D_2O were found to be equal, within experimental error (*10*). The purpose of this study was to have a better understanding of the transition state stabilization by MR. Through these studies, the difference between the ground and transition states of the substrate was determined to be $\Delta G^{\ddagger}$ = -26 kcal/mol.

Thermodynamic data reported from St. Maurice and Bearne (2002) include substrate binding enthalpy and entropy changes (ΔH = -8.9 ± 0.8 kcal/mol and $T\Delta S$ = -4.8 ± 0.8 kcal/mol), activation barrier for conversion of bound substrate to bound product ($\Delta H^{\ddagger}$ = 15.4 ± 0.4 kcal/mol, $T\Delta S^{\ddagger}$ = 2.0 ± 0.1 kcal/mol), and transition state stabilization (ΔH_{tx} = -22.9 ± 0.8 kcal/mol and $T\Delta S_{tx}$ = 1.8 ± 0.8 kcal/mol) (*4*). All of this data pertains to the catalysis of (*R*)-mandelate by MR at 25 °C, pH 7.5. The magnitude of both the ΔH_{tx} and $T\Delta S^{\ddagger}$ indicates that proficiency of the enzyme is attained by reduction

in enthalpy, although there are entropic contributions as well. The signs of $\Delta H^{\ddagger}$ and $T\Delta S^{\ddagger}$ indicate that mandelate racemase stabilizes the transition state for proton abstraction by reducing the enthalpy of activation and increasing entropy of activation in relation to the nonenzymatic reaction (*4*). St. Maurice and Bearne (2002) also suggest that MR has reached it limit of evolution due to it being close to the theoretical limit of catalytic proficiency, i.e. close to diffusion limitation (*4*).

3.1.4 Other Substrates and Inhibitors

3.1.4.1 Studies investigating mandelate analogs

When the native enzyme was first purified and characterized, the Kenyon group (1970) also investigated the reactivity using a few mandelate analogs. In these studies the analogs were substituted in the 4-position of the phenyl group (*5*). It was found that 4-bromo and 4-chloro gave an enhanced V_{max} value, through stabilization of the negative charge on the α-C atom by these electron-withdrawing groups. Conversely, 4-hydroxy and 4-methoxy, both electron-donating groups, gave a decreased V_{max} value, due to their electron donating abilities (Table 3-2-Data adapted from (*5*)).

This data expressed quantitatively can be used in predictions of other reaction rates and equilibria, as well as give a better understanding of the reaction mechanism. In this case, the relationship between rates and equilibria can then be measured by the Hammett equation, with a

measured kinetic σ value of approximately +2 (*5*). This involves reactions containing substituted phenyl groups.

Table 3-2. Kinetic data for mandelate analogs.

Analog	V_{max} (μmoles/min)	K_M ($M \times 10^5$)
4-Bromo	124 ± 24.8	25.6 ± 5.1
4-Chloro	107.5 ± 21.5	10.0 ± 2.0
4-Hydrogen	33 ± 6.6	9.3 ± 1.9
4-Hydroxy	14.8 ± 3.0	29 ± 5.8
4-Methoxy	5.5 ± 1.1	33 ± 6.6

The Kenyon group (1970) also found analogs that served as partial inhibitors such as those found in Table 3-3. This inhibition, however, only occurred when incubated with enzyme prior to mandelate addition. The inhibition was reversible and no clear pattern for molecular structures

Table 3-3. Mandelic acid analogs that serve as reversible inhibitors of MR.

Analog Tested	% Inhibition
Phenyloxyacetic acid	83
Phenylmercaptoacetic acid	83
Phenylacetic acid	74
DL-β-Phenyllactic acid	65
DL-Atrolacetic acid	46
Diphenylacetic acid	19
Benzoic acid	7

leading to inhibition was found. In later work also by the Kenyon group (1974) α-phenylglycidate (Figure 3-1) was determined to cause complete and irreversible inhibition (K_I = 1.92 mM) of MR activity (*13*). The enzyme can be protected from inactivation by both of the mandelate enantiomers. Benzilate (Figure 3-1) acts as a competitive inhibitor (K_I = .67 mM) with both of its phenyl groups mimicking the simultaneous binding of the (*R*)- and (*S*)-mandelate phenyl groups (*14*).

α-hydroxy-benzylphosphonate (α-HBP) (Figure 3-1) has been identified as a potent, competitive inhibitor (K_I = 4.7 μM), acting as an

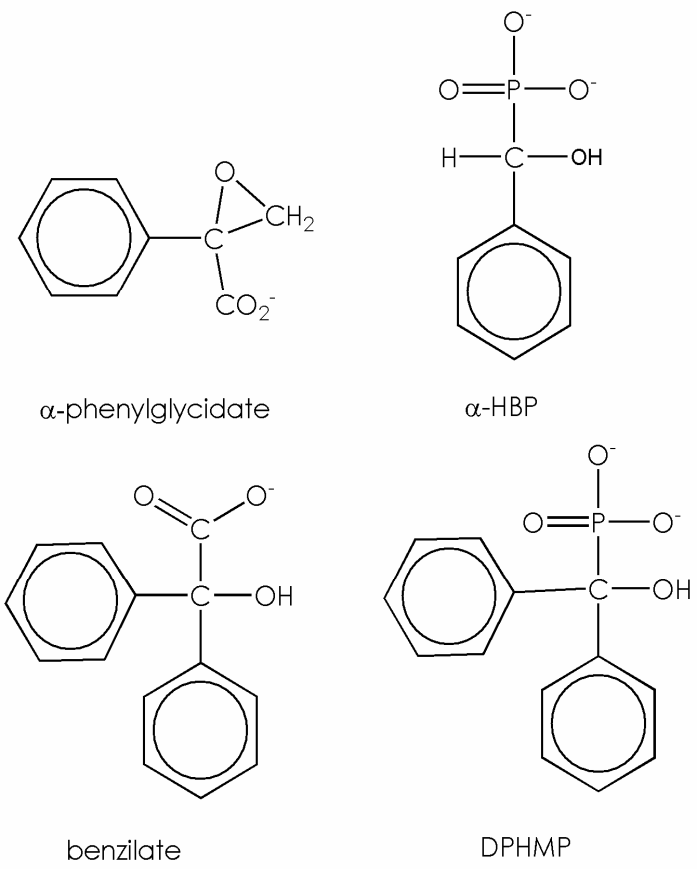


Figure 3-1. Mandelate racemase inhibitors.

analog to the reaction intermediate. Another inhibitor that has been studied is 1,1-diphenyl-1-hydroxymethylphosphonate (Figure 3-1). The two phenyl groups of this compound are suggested to bind in both the (*R*)- and (*S*)-pockets, leading to strong competitive inhibition with a K_I value of 1.41± 0.09 mM (*14*).

3.1.4 Mechanism for MR

Mandelate racemase catalyzes the abstraction of an α-proton from a carbon adjacent to a carboxylic acid group (Figure 3-2). This is a two-base mechanism where two general base/general acid catalysts transfer protons to and from the α-carbon of the substrate and product enantiomers, forming an enolic intermediate. Through site-directed mutagenesis of key residues and careful analysis of the crystallographic structure, a great deal of evidence for this mechanism has been collected (*15-18*).

A high-resolution tertiary structure identified the two catalytic bases as K166 and H297. Catalysis occurs via a two-base mechanism with H297 responsible for abstracting the α-proton from (*R*)-mandelate and K166 responsible for abstracting the α-proton from (*S*)-mandelate. The general acid catalyst is E317 for both enantiomers, as it stabilizes the enolic intermediate (*17*). A variety of residues are involved in hydrophobic interactions binding the substrate to the active site.

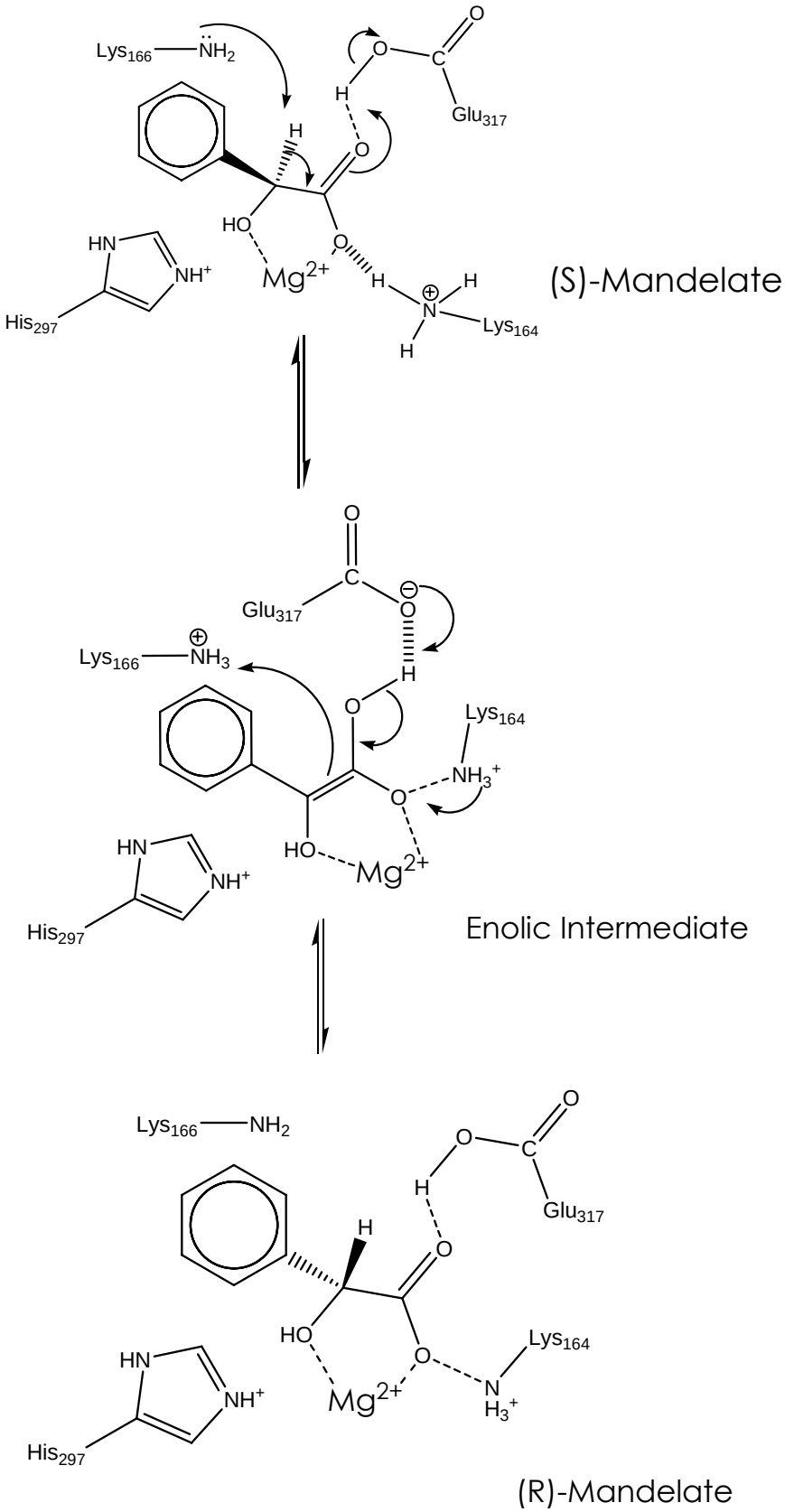


Figure 3-2. Mechanism for reaction catalyzed by mandelate racemase.

The Gerlt group has done extensive studies on the MR mechanism through analyzing mutations of the predicted active site residues (*15*, *17*,*18*). A H297N mutation does not lead to catalysis of the racemization of either enantiomer, it does however catalyze the facile exchange of the α-proton of the (*S*)- but not (*R*)-enantiomer. A K166R mutation retains a low level of activity and catalyzes conversion of the (*R*)- but not (*S*)-enantiomer (*18*). Site-directed mutagenesis was also used to explore the general acid catalyst E317, the Gerlt group found that the E317Q mutant catalyzes the reaction at a very low rate (*17*).

Therefore, it has been concluded that there are three basic steps for the mechanism of the MR reaction. The first step is the catalytic protonation of the carboxylate group in the mandelate substrate by the hydrogen bonded residues Lys164 and Glu317. The next step involves the configuration inversion through the deprotonation of mandelic acid by Lys166 and the back donation by His297. In the final step the catalytic proton of the carboxylic group is back-transferred to the Lys164/Glu317.

One problem here is that at pH 7.5 where MR is active, the ε-amino group of lysine would be protonated, however the mechanism requires free basic Lys for removal of the α-proton, unless the pK_a of the Lys has been shifted to a lower value than normal. Kenyon and Gerlt detemined that the pK of Lys166 is 6, and through the x-ray crytal structure of the enzyme revealed an ammonium salt group of a second Lys (Lys164) (*19*).

The ammonium salt group of Lys164 is close enough to the Lys166 to produce a large electrostatic effect (*20*).

The active site of mandelate racemase is composed of a positive electrostatic environment, however the pK_a of the imidazolium group is not affected. The Gerlt group has determined that hydrogen-bonding interaction between Asp270 and His297 is important for the appropriate positioning of the imidazole group of His297 for the facile abstraction of the α-proton of (*R*)-mandelate. Gerlt's group also suggests the Asp270 is responsible for "insulating" His297 from the positive electrostatic environment that appears to lower the pKa of the ε-ammonium group of Lys166 from its normal value (*17*). Mutating Asp270 to an Asn and thus preventing the hydrogen bond from forming causes the side chain of His297 to move about 0.5 Å away from the Asn and toward the substrate (*2*). There is also an observed reduction in the k_{cat} value by a factor of 10^4 for both directions of the reaction. As with the H297N mutant, the D270N mutant catalyzed the exchange of the α-proton of the (*S*)- but not (*R*)-enantiomer. Results from this experiment suggest that D270 and H297 operate as a catalytic dyad (*17*).

3.1.5 Metal Requirement for MR

Mandelate racemase requires a divalent metal ion for activity. It has been found that the metal ion stabilizes the enolate anion intermediate (*12*). While the preferred ion is Mg^{2+}, other divalent metal

ions Co^{2+}, Ni^{2+}, Mn^{2+}, and Fe^{2+} also stabilize the enolate intermediate resulting in an active enzyme. Divalent metal ions that are unable to restore activity include: Ca^{2+}, Sr^{2+}, Ba^{2+}, Cu^{2+}, Zn^{2+}, and Cd^{2+}. Although the need for a metal ion seems to lower the probability of ever finding such an enzyme in the solar system, magnesium is ubiquitous on rocky planets and moons and thus the presence of magnesium ion only requires water. In comparison, amino acid racemases require other cofactors mostly pyridoxal-5'-phosphate for their activity.

3.1.6 Aim 1: Clone, Overexpress, Purify, and Characterize Mandelate Racemase from *Pseudomonas putida*

The purpose of this Aim is to isolate and clone the mandelate racemase gene from *P. putida*, to transform this gene into *E. coli* (JM105) and overexpress the protein. As a hexa-histidine-tag has been added to the N-terminus of the gene, purification via immobilized metal affinity chromatography is straightforward. Characterization of the enzyme is performed with activity (via polarimetry), and compatibility with different metal ions of activity. One of the major goals of this aim is to determine activation parameter (E_a, $\Delta H^{\neq}$, $\Delta S^{\neq}$, $\Delta G^{\neq}$) at standard conditions.

3.2 Materials and Methods

3.2.1 Materials

(*R*)-mandelic acid was a kind gift from BASF (Ludwigshafen Germany); (*S*)-mandelic acid, citric acid, ammonium sulfate, Anti-foam 204, and Tris(hydroxymethyl)-aminomethane were obtained from Sigma-Aldrich (St. Louis, Missouri), $MgCl_2$ was from Fisher Chemical (Fair Lawn, New Jersey), Luria Bertani (LB) broth, Aerosol OT, ammonium acetate, heptane and sodium phosphate from EM Science (Gibbstown, New Jersey), ampicillin from Shelton Scientific , ammonium formate and 1-hexanol from Alfa Aesar (Pelham, New Hampshire).

3.2.2 MR Expression/Purification from *Pseudomonas putida*

Pseudomonas putida ATCC 12633 was grown in a medium containing both glucose and mandelate as carbon sources at 30°C as described in literature (*5*). The cells were harvested by centrifugation at 27,000 g for 45 minutes, one hour after the previously described color change from grey to yellow. The cell pellets were lysed using sonication, six times for 2 minutes each, while on ice. The cell lysate was heat-treated at 55°C for 2 hours and then centrifuged at 27,000 g for 45 minutes. Next, the supernatant was dialyzed against 40% ammonium sulfate overnight at a pH of 7.5 and 4°C. After centrifugation at 27,000 g for 30 minutes, the protein was purified using ion exchange chromatography. A Q-sepharose 20 mL column (flowrate onto column 1 mL/min, elution flowrate 3 mL/min,

1 M NaCl gradient 0% to 100) connected to an ÄKTA FPLC system was used with the enzyme eluted from the column with a linear sodium chloride gradient.

3.2.3 General DNA Techniques

Extraction of plasmid DNA from *E. coli* was prepared using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA). DNA was purified using the DNA Clean & Concentrator kit (Zymo Research, Orange, CA), according to manufacturers instructions. Agarose gel electrophoresis was performed using a 1.2% TAE agarose gel according to Sambrook (2001) (*21*) and were stained with GelStar® Nucleic Acid Gel Stain (BMA, Rockland, ME) according to manufacturers instructions. Fragments of DNA were isolated from the agarose gel using the QIAquick Gel Extraction Kit or Qiaex II (Qiagen, Valencia, CA) with a 1 kb ladder (New England Biolabs, Beverly, MA) for molecular weight standards. Restriction enzymes were used according to the suppliers *Nde*I, *Bam*HI (New England Biolabs, Beverly, MA), *Eco*RI, *Hind*III, and *Dpn*I (Fermentas, Hanover, MD). Gene fragments were ligated using either DNA Ligation Kit (Novagen, Madison, WI) or Quick Ligation Kit (New England Biolabs, Beverly, MA), according to the suppliers' instructions. Competent *E. coli* XL1-Blue cells were prepared and plasmids transformed according to the manufacturers instructions.

The DNA polymerase chain reactions (PCRs) were carried out in a programmable Mastercycler® gradient thermocycler (Eppendorf,

Westbury, NY). A "hotstart" PCR program was used for gene amplification, with DNA predenatured at 98°C for 5 minutes, subsequently *Pfu* DNA polymerase enzyme was added according to manufacturers instructions. The following sequence was used for 25 cycles: 95°C for 30 seconds, 5°C less than T_M for 30 seconds, and 72°C for 1 minute 15 seconds. An elongation cycle was added at the end at 72 °C for 10 minutes. All PCR's were run using this method unless otherwise noted.

3.2.4 Cloning, fermentation, and expression

Pseudomonas putida genomic DNA was prepared according to Murphy (2002) (*22*). The gene for mandelate racemase was extracted from the genomic DNA via PCR using primers designed for this purpose (Table 3-4). The gene was cloned into the high copy plasmid pUC18 using the restriction sites *EcoR*I and *Hind*III. The resulting recombinant plasmid was then transformed into the *E. coli* strain XL1Blue. Plasmid DNA from positive clones was then prepared and sequenced at the FAME Sequencing Center, Emory University, Atlanta, GA.

Table 3-4. Primers used in cloning of the MR gene.

	Forward primer (5')	Reverse primer (3')
pUC18 primers	gcgc***gaattc***atgagtgaa gtactgattaccggcc	gcgc***aagctt***ttacaccagatatttcccgatttctttctc
pkk223 primers (His-Tag primers)	gcgc***gaattc***atgagtgaa gtactgattaccggcctg	gcgc***aagctt***ttacaccag atatttcccgatttcttt

Legend: Italics denotes restriction sites *EcoR*I and *Hind*III.

The gene was then cloned into the expression vector pkk223-3 using a primer that included a N-terminal hexa-histidine tag, along with *EcoR*I and *Hind*III restriction sites. The construct was transformed into the *E. coli* strain JM105, cells that were made competent according to manufacturers instructions. The strain was grown at 37°C in LB broth containing 100 μg/ml ampicillin. For expression of protein, one colony was picked and grown to mid-log phase (OD_{600nm} = 0.6). Protein expression was induced with 0.1 mM isopropyl-β-thiogalactoside (IPTG), an additional 100 μg/ml ampicillin was added, to maintain selection pressure, and 20 μL of antifoam was added. Four hours after induction the cells were harvested by centrifugation for 30 minutes at 27,000 g and frozen at –80 °C.

3.2.5 Purification of Recombinant Protein

The cell pellet was thawed and resuspended in lysis buffer consisting of 50 mM Tris buffer (pH 7.5), 1mM $MgCl_2$ at a concentration of 5 mL per gram of wet cell pellet. This suspension was then sonicated six times for 2 minutes each to lyse the cells, allowing cooling on ice between each burst, and then centrifuged at 27,000 g to remove cell fragments. Employing the hexa-histidine tag, the enzyme was purified from cell lysate using nickel-nitrilotriacetic acid (Ni-NTA) metal affinity chromatography following Qiagen's protocol (Qiagen, Valencia, California). To verify the imidazole from the hexa-histidine tag does not have an effect on the

enzymatic activity, we added 0.25 M imidazole in a solution of 100 mM mandelate, pH 7.5, T = 22 °C and measured the change in optical rotation over an eighteen hour period.

3.2.6 Apo-Enzyme

The apo-enzyme was produced by dialyzing the purified enzyme against 1 mM EDTA for fourteen hours at pH 8.0 and and 4°C to remove the bound Mg^{2+} ion (*12*). Various other divalent metal ions were added as cofactor with concentrations up to 100 mM.

3.2.7 Enzyme Assay

Purified enzyme was assayed using a polarimetric assay. All polarimetric measurements were taken on a Jasco P-1010 at 390 nm using a cuvette of 10 cm pathlength with a total volume of 8 mL. The reading of optical rotation versus time was recorded at 24°C and pH 7.5 with 100 mM mandelate substrate in an aqueous solution of 50 mM Tris buffer, and 1 mM $MgCl_2$ to ensure adequate cofactor presence. The specific activity was then calculated using equation 3-1.

$$SA = \frac{[S]}{\alpha_S} * \frac{1}{2} * \frac{1}{[E]} * V_S * V_E * \frac{\Delta \alpha}{time} \qquad 3\text{-}1$$

where [S] is the substrate concentration, α_s is the optical rotation of the substrate, [E] is the enzyme concentration, V_S is the volume of substrate

solution, V_E is the volume of the enzyme added to reaction, $\Delta\alpha$ is the change in optical rotation.

3.2.8 Polarimeter Measurements

The optical rotation of 13.1 mM mandelate (50 mM Tris buffer, pH 7.5, 23°C) was measured versus wavelength over various wavelengths and the measured optical rotation was converted to specific optical rotation. The specific optical rotation, [α], is defined as the observed angle of optical rotation when plane-polarized light is passed through a sample (1 g/100 mL) with the path length of 1 dm, this is represented by equation 3-2:

$$[\alpha]_{\lambda}^{T} = \frac{\alpha}{l * c} \qquad 3\text{-}2$$

where [α] is the specific optical rotation, T is the temperature of measurement, λ is the wavelength of the measurement, a is the measured optical rotation, l is the pathlength in dm, and c is the concentration of the solution (g/100 mL).

3.3 Results

Since a polarimetric assay is used for this enzyme, initial studies were performed to determine the change in optical rotation, α, as a function of mandelate concentration, change in wavelength, and temperature. The averages of multiple optical rotation readings at each condition were charted, respectively (Figures 3-3, 3-4, 3-5).

We found that α decreases monotonically with increasing wavelength λ. While measuring at the smallest wavelength produced by the tungsten lamp (340 nm) would provide the largest signal-to-noise ratio, we picked 390 nm as the wavelength to measure enzyme assays. A wavelength of 390 nm results in a large signal while staying away from the instruments limitations. We found that optical rotation α is strictly proportional to mandelate concentration, with a change of a specific rotation of 14.4° ± 1.8° /mM. Solubility of mandelate in aqueous buffer is around 300 mM, at a pH of 7.5. From Figure 3-3 we determined that the optimal mandelate concentrations for assays is100mM.

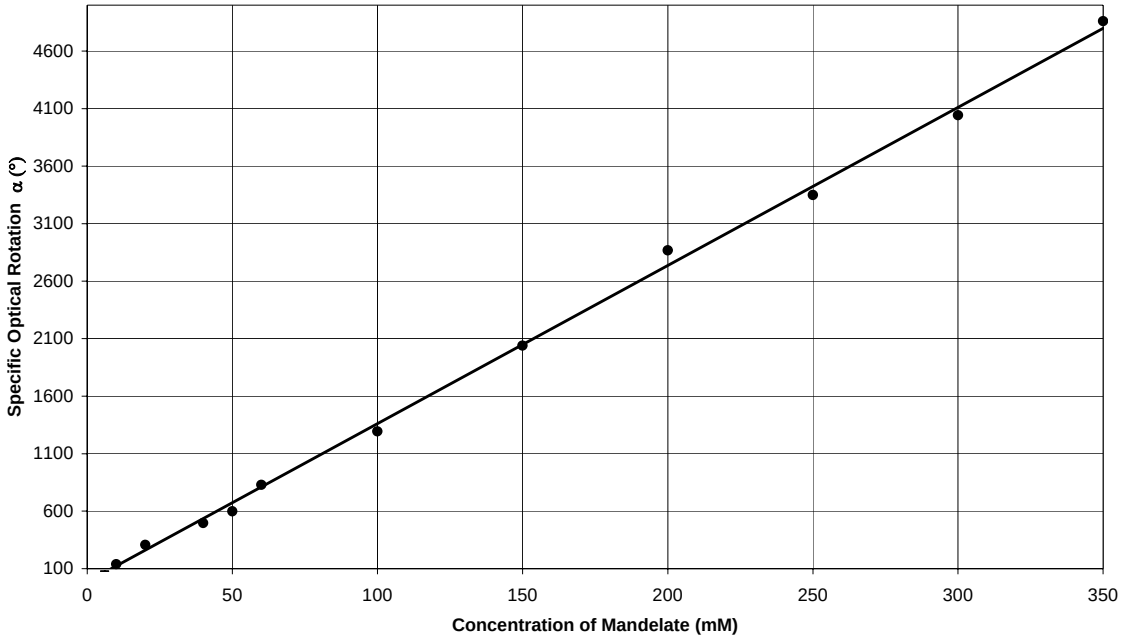


Figure 3-3. The relationship between optical rotation and concentration using mandelate. Conditions: pH = 7.5, T = 22 °C, λ = 589 nm.

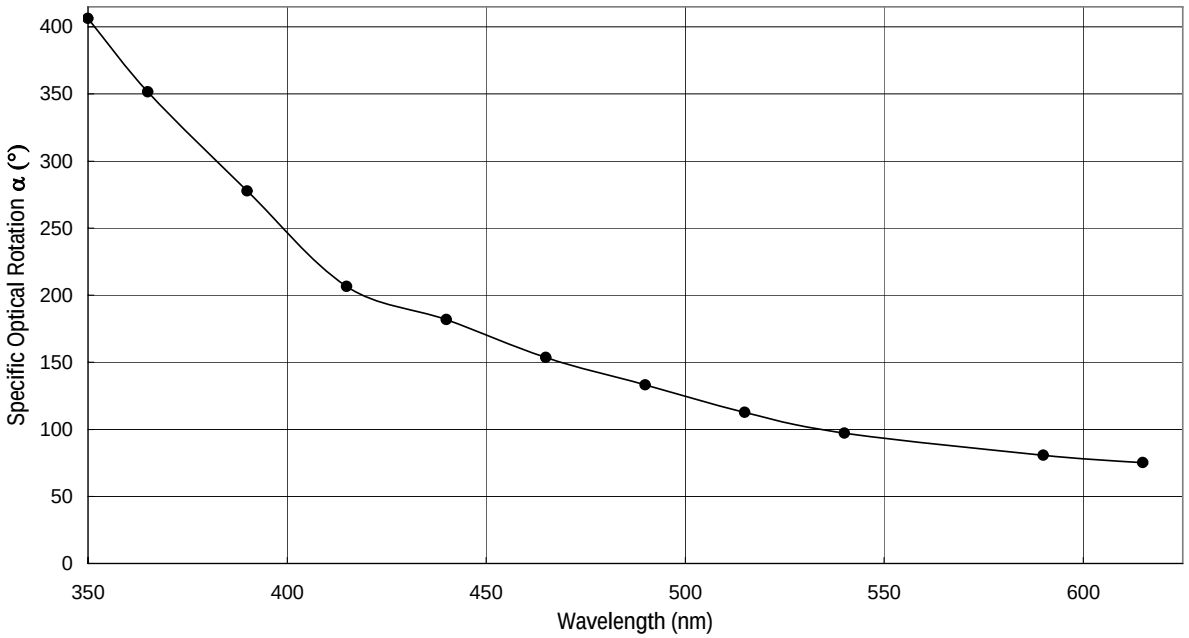


Figure 3-4. Specific optical rotation of mandelate versus wavelength. Conditions: [mandelate] = 13.1 mM, pH = 7.5, T = 22 °C.

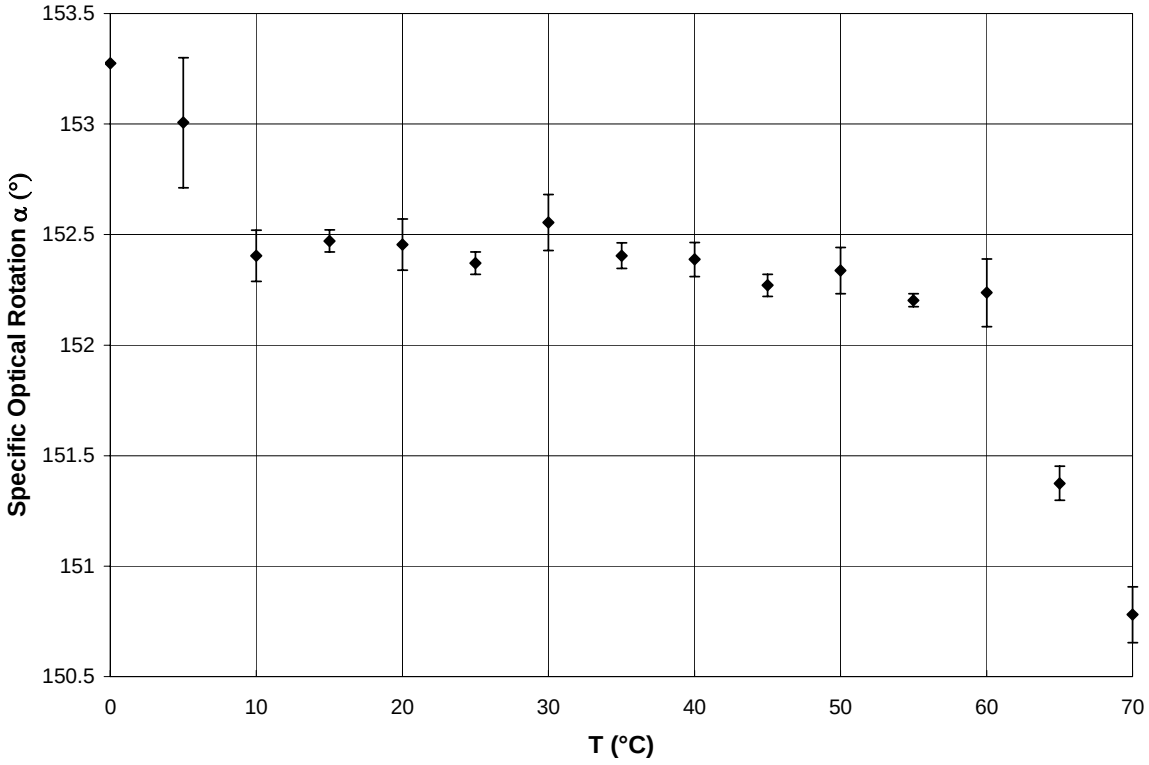


Figure 3-5. Optical rotation versus temperature. Conditions: [Mandelate] = 13.1 mM, λ = 589 nm, pH = 7.5.

3.3.1 MR cloning, overexpression, and protein purification

Mandelate racemase was successfully purified from *P. putida* cell lysate using several steps. The first step of purification was heat treatment, which resulted in 2.5 fold purification. The next step was a 40% ammonium sulfate precipitation, followed by Q-sepharose column with a purification factor of 49 (Table 3-5) with a purity of 82% as determined by densiometry.

The gene for MR was successfully cloned from genomic DNA of *P. putida* pUC18 for storage and then cloned into pkk223-3 for overexpression. This recombinant plasmid was transformed into JM105 cells (*E. coli*) and overexpressed (37°C with 0.1 mM IPTG to OD_{600nm} = 1.2) to 4.5% of cell protein. The protein has been purified taking advantage of a N-terminal 6X histidine tag using a nickel-NTA purification column. Purity of the enzyme determined from densiometry was 94% with a purification factor of 23 (about 80 mg per 1 L culture) (Figure 3-6). There was a slight background reaction between the substrate and imidazole of 0.002 ± 0.0001 U/mg.

Table 3-5. Purification table for both non his-tag MR and his-tag MR.					
Sample	Volume (mL)	Total Activity (U)	Total Protein (mg)	Specific Activity (U/mg)	Purification Factor
Non His-Tag Purification					
Lysate	24	2.4	1170	2	1
Heat Treated	21	2.8	996	4	2.5
40 % $(NH_3)_2SO_4$	12	3.0	318	9	4.5
Q-sepharose	7	3875	569	97	49
His-Tag Purification					
Lysate	20	15800	1420	11	1
1st elution	5	11000	46	255	23
2nd elution	5	2000	2825	7	0.6
3rd elution	5	35	11	1.3	0.1

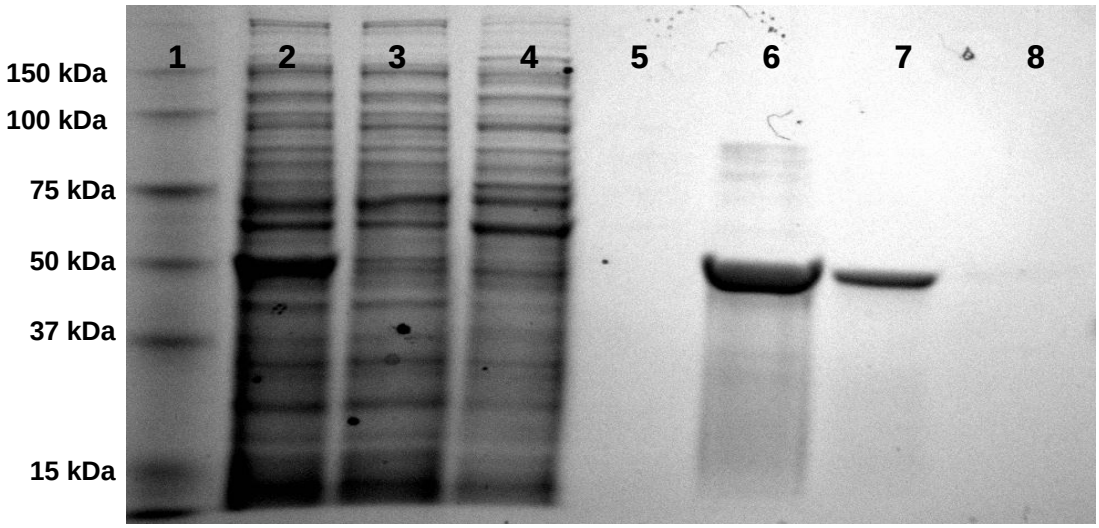


Figure 3-6. SDS-PAGE gel of purified his-tag MR. Lane 1: Molecular weight markers, Lane 2: Lysate (loaded 30 μg), Lane 3: Flow through column (loaded 30 μg), Lane 4: First wash (loaded 30 μg), Lane 5: Second wash (loaded 15 μg), Lane 6: First elution(loaded 15 μg); Lane 7: Second elution (loaded 15 μg), Lane 8: Third elution (loaded 15 μg)

3.3.2 Apo-Enzyme Studies

Apo-enzyme was prepared to determine which divalent metal ions besides Mg^{2+} in the metal-binding site result in active enzyme. Mg^{2+}, Co^{2+}, Ni^{2+}, and Mn^{2+} all resulted in active enzyme, as reported in literature. Other metal ions (Li^{+}, Ca^{2+}, Cu^{2+}, Zn^{2+}, Sr^{2+}, Ru^{2+}, Pd^{2+}, $Cd^{2+,}$ Ba^{2+}, Hg^{2+}, and Pb^{2+}) were tested with no subsequent activity measured (Table 3-6). These results suggest that divalent metal ions with a radius larger than that of Mg^{2+} (0.86 Å) do not result in an active enzyme. As expected, activity was retained after adding $MgCl_2$ back to the apo-enzyme. Decreasing activity was found with decreasing Mg^{2+}. When adding 100 mM of metal ions other than Mg^{2+}, followed by the addition of 10 mM Mg^{2+} ion, only the enzyme with prior addition of Li^{+} and Ru^{2+} showed activity when reconstituted with Mg^{2+}. The addition of other metals resulted in inactive enzyme. Therefore, we were able to show that of the catalytically inactive metal ions, Li^{+} and Ru^{2+} do not associate with the metal-binding site of the enzyme, while the other metal ions apparently bind and lead to inactive enzyme.

Table 3-6. Metals tested with MR.

Metal	Specific Actiivty (U/mg)
Mg^{2+}	419
Mn^{2+}	50
Co^{2+}	288
Ni^{2+}	241
Cu^{2+}	0
Zn^{2+}	0
Li^{+}	0
Ca^{2+}	0
Pd^{2+}	0
Ru^{2+}	0
Pb^{2+}	0
Hg^{2+}	0

We determined the concentration of magnesium chloride that needs to be added (±12%) in order to restore activity. As little as 1 mM results in 85% of the initial activity, measured at 13.1 mM, while 10 mM and 100 mM yield 96% (± 17%), and 113% (± 16%), respectively.

3.3.3 Kinetic Studies

Using the polarimetric assay, the specific activity of purified enzyme was measured at 24°C, pH 7.5 to be 320 ± 35 U/mg. Michaelis-Menten, Lineweaver-Burk, and Eadie Hofstee plots were constructed from the kinetic data (Figures 3-7, 3-8 and 3-9, respectively). We were unable to measure kinetics below about 5 mM of substrate because of lack of sensitivity in the polarimetric assay. The K_M and v_{max} values were calculated and can be found in Table 3-7.

Table 3-7. MR kinetic data using two different methods.

Method	K_M (mM)	v_{max} (μM/s)	k_{cat} (s^{-1})	k_{cat}/K_M (M^{-1} s^{-1})
Lineweaver-Burk	10.1 ± 1.1	(5.6 ± 0.5)*10^3	1880 ± 160	1.8*10^5
Eadie Hofstee	10.2 ± 0.9	(5.7 ± 0.4)*10^3	1870 ± 130	1.8*10^5

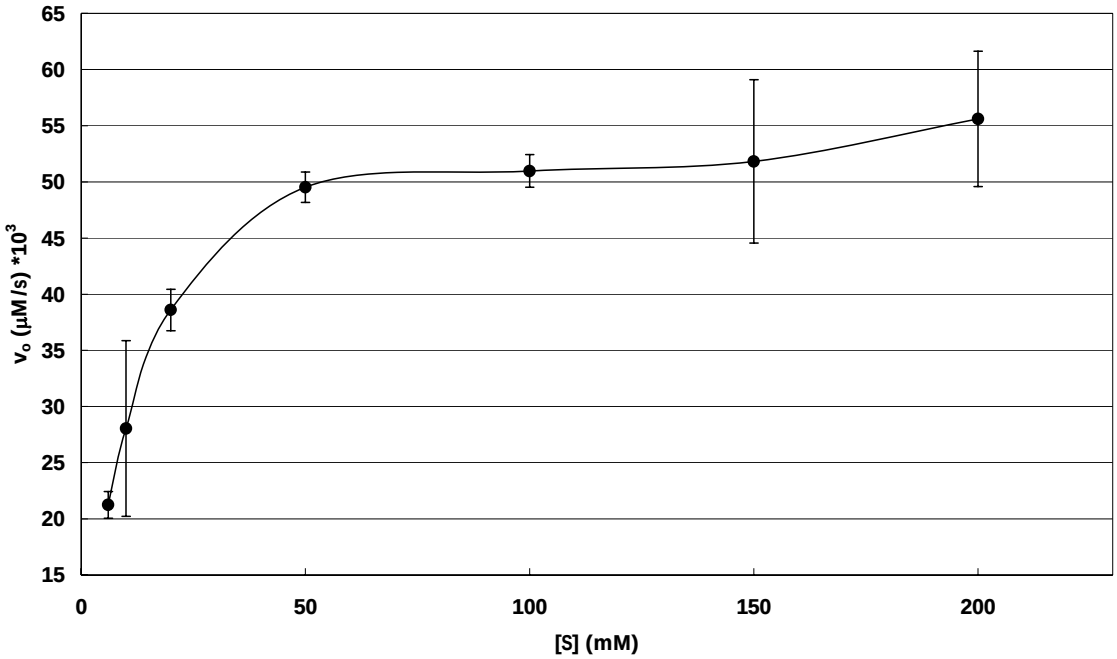


Figure 3-7. Michaelis Menten plot for MR kinetics. T = 23 °C; pH 7.5; and [E]= 0.3 μM.

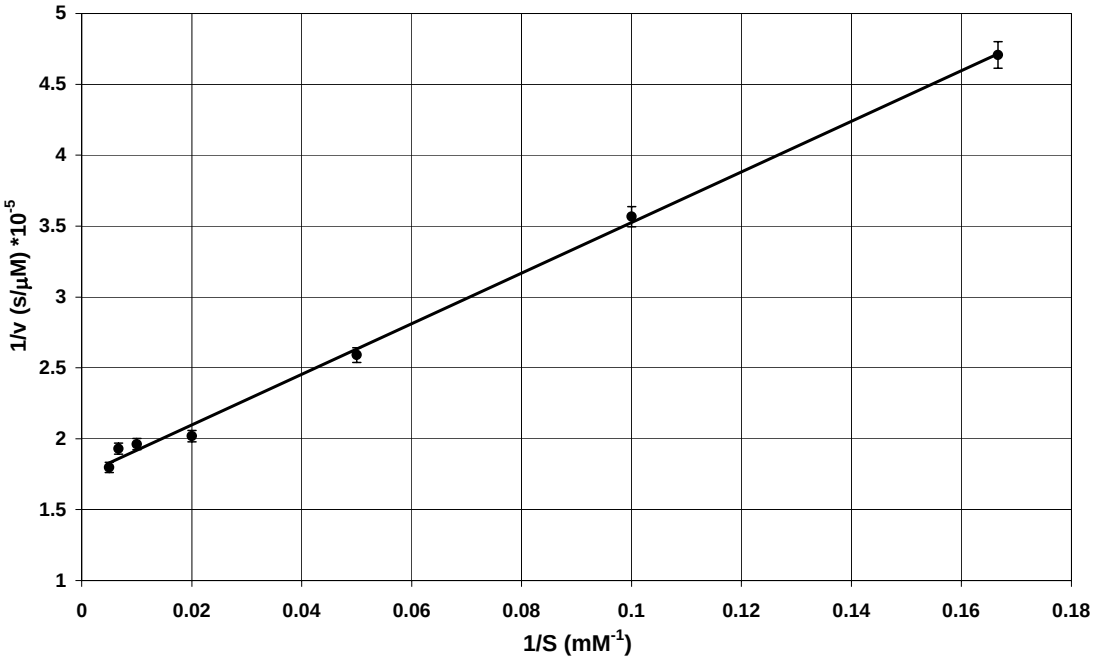


Figure 3-8. Lineweaver-Burk plot for MR kinetics. T = 23 °C; pH 7.5; and [E]= 0.3 μM.

Kinetic data determined from both Lineweaver-Burk and Eadie-Hofstee plots with polarimetry were in agreement with a K_M values of 10 ± 1 mM and a maximum velocity v_{max} of about $5.7*10^3$ mol/(s·L). From the specific activity v_{max}/[E] (in U/mg = μmol/(min·mg protein)), with the molecular mass of the enzyme of 347 kDa, and the knowledge that MR is composed of eight identical subunits (*2*), we obtain a catalytic rate constant k_{cat} of 1870 ± 145 s^{-1}.

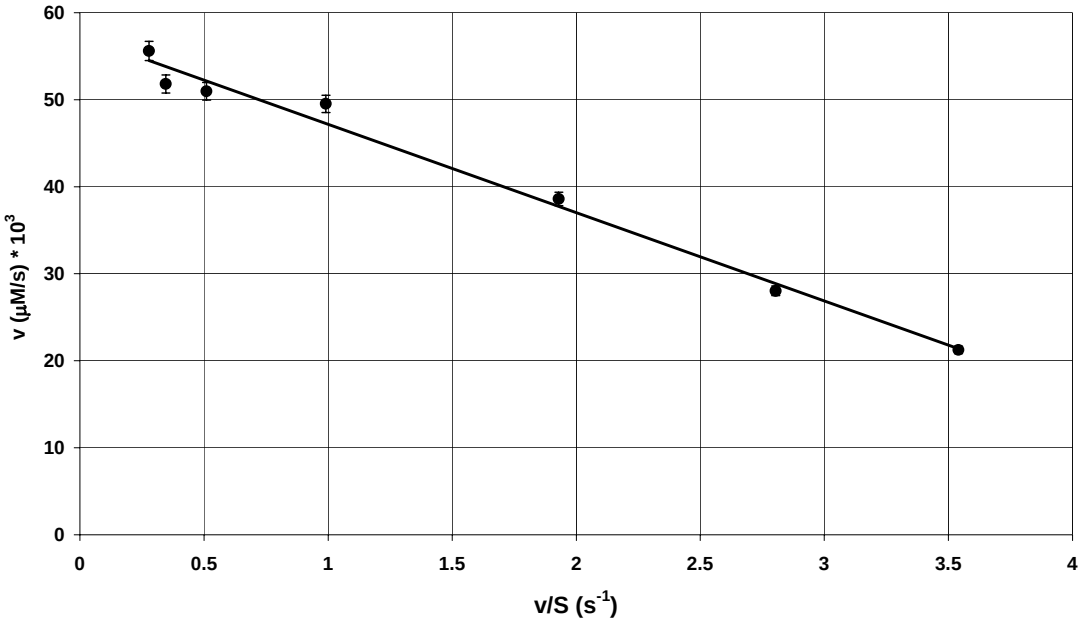


Figure 3-9. Eadie-Hofstee plot for MR kinetics. T = 23 °C; pH 7.5; and [E]= 0.3 μM.

3.3.4 Temperature Studies

We measured the specific activity of MR using 100 mM mandelate in 50mM Tris buffer at pH 7.5 and the temperature as a parameter and obtained data between 0°C and 70°C with an optimum of reactivity at 50°C (Figure 3-10). As the buffer freezes at -3°C, we were able to collect a data point at 0°C. We plotted ln k_{cat} versus 1/T (Arrhenius plot; Figure 3-11) and obtained the activation energy E_a from the slope ($-E_a/R$); the plot has a correlation coefficient of 0.98. By plotting ln k_{cat}/T over 1/T we obtained the Eyring plot (Figure 3-12), and determined the enthalpy of activation ($\Delta H^{\neq}$) from the slope ($-\Delta H^{\neq}/R$) and the entropy of activation $\Delta S^{\neq}$ from the

intercept according to $y(x = 0) = \ln (k_B/h) + (\Delta S^{\neq}/R)$ (Table 3-8); the plot has a correlation coefficient of 0.94. The Gibbs free enthalpy of activation was calculated from $\Delta G^{\neq} = \Delta H^{\neq} - T \cdot \Delta S^{\neq}$. The activation energy was determined for the deactivation temperature data to be $E_{a,d} = 62 \pm 5.3$ kJ/mol (Figure 3-13).

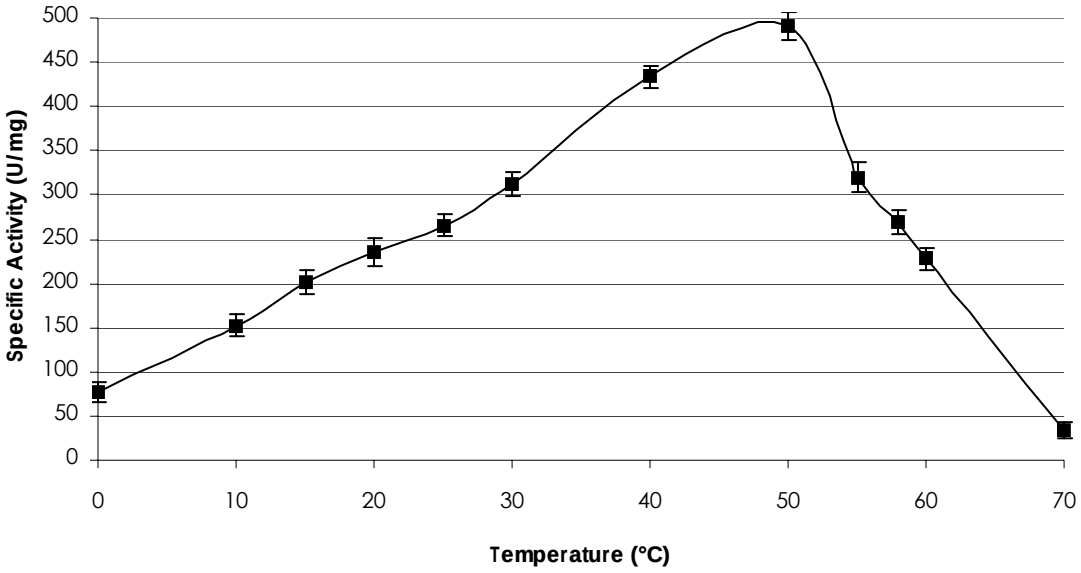


Figure 3-10. Specific activity of MR over a temperature range of 0°C to 70°C.

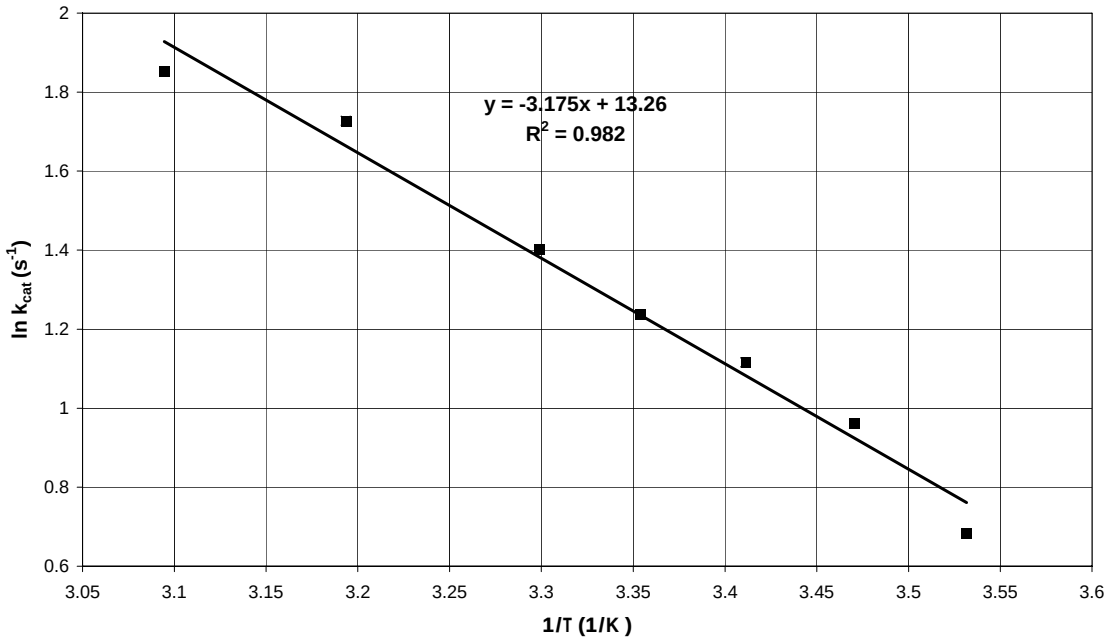


Figure 3-11. Arrhenius plot of MR in Tris buffer, pH 7.5. [E]= 0.3 μM.

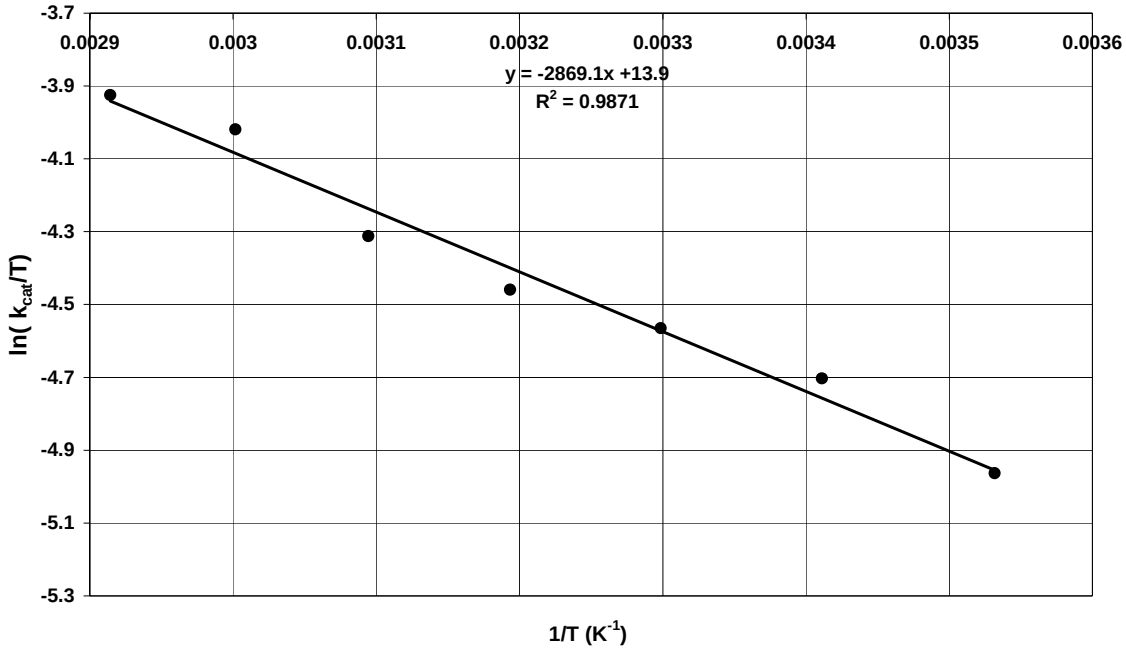


Figure 3-12. Eyring plot for MR in Tris buffer, pH 7.5, [E]= 0.3 μM.

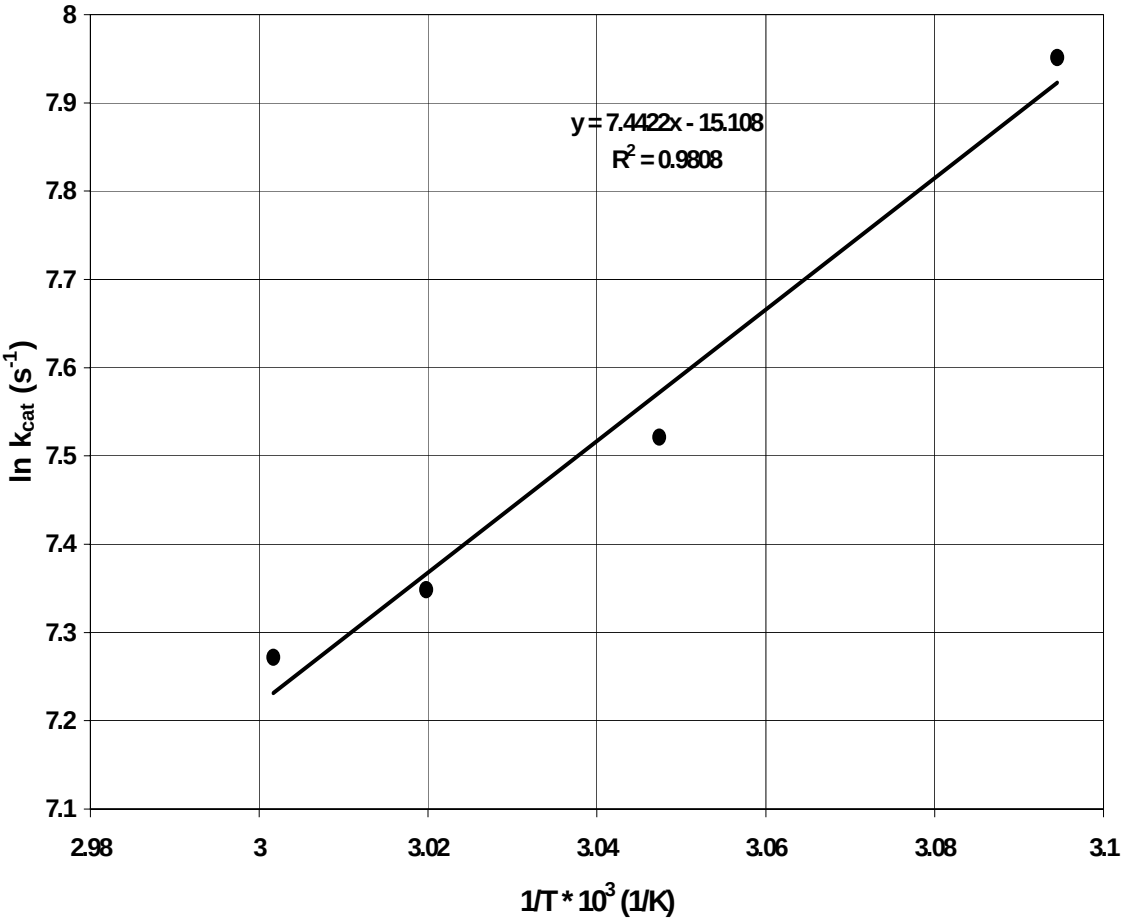


Figure 3-13. Arrhenius plot for deactivation of MR in Tris buffer, pH 7.5. [E]= 0.3 μM.

3.4

Table 3-8. Activation parameters for MR in aqueous buffer, pH 7.5, T = 25 °C. [a]St. Maurice et. al (2002) pH 7.5, T = 25 °C. [b]Bearne et. al. (1997) pH 7.5, T = 25 °C

	E_a (kJ/mol)	$\Delta H^{\neq}$ (kJ/mol)	$\Delta S^{\neq}$ (J/mol K)	$\Delta G^{\neq}$ (kJ/mol)
MR HEPES (lit)[a]	NR	65 ± 2	8.4 ± 0.4	56 ± 2
MR PO_4 (lit)[b]	NR	134 ± 4	-39 ± 2	147 ± 9
MR Tris (this work)	26 ± 1.3	24 ± 1.2	-118 ± 20	54 ± 2

3.4 Discussion

It was important as a first step towards obtaining accurate measurements of the activity of MR using polarimetry, to determine the optimal conditions under which to assay the enzyme. For this reason we began our studies by determining which parameters have a significant effect on the optical rotation of mandelate. As an initial check of the instrument, we demonstrated a linear correlation between mandelate concentration and optical rotation (Figure 3-3). We then followed this by determining the specific optical rotation per mM concentration, of mandelate at various temperature and pH values. Having an accurate value for the specific optical rotation under various conditions is important, as this parameter is used to determine the amount of mandelate consumed in the reaction and thus the activity of the enzyme. No clear trend could be found to relate changes in pH to optical rotation. There is a slight temperature difference at the extremes, hence it is best to take measurements only in the consistent temperature range (between 10°C and 60°C). This temperature data allowed us to convert changes in optical rotation over time accurately back to change in mandelate concentration at any given temperature. It should be noted that the fluctuation at the extremes is within the range of error for the polarimeter. As expected we found a linear correlation between concentration and

optical rotation (Figure 3-3). We determined a specific rotation $[\alpha]_{390}^{24}$ = 175° for mandelate, pH 7.5.

The specific activities we obtained for our his-tagged enzyme were higher than previously published values for the wild-type. In our case, we determine a specific activity of 320 ± 35 U/mg protein (24°C, pH 7.5) compared to the published value of 195 U/mg for the untagged protein (*5*). Our k_{cat} value is also higher than previously reported values for k_{cat} for the untagged protein, which is 500 s^{-1} (*5*) compared to our value of 1875 ± 145 s^{-1}. These differences in specific activity and k_{cat} values may either be due to changes in the properties of the enzyme caused by the his-tag or may be due to the higher level of purity provided by the his-tag. On the other hand, we calculated a K_M value of 10.1 ± 1.1 mM, which is an order of magnitude higher than previously published values which range from 0.34 mM (*17*) to 0.81 mM (*5*). Such a big difference however, cannot be easily explained by the his-tag, and upon closer inspection demonstrates one of the limitations that we face using a polarimetric assay, namely a lack of sensitivity, which in this case prevents us from assaying the enzyme accurately at substrate concentrations below 5 mM (*9*). This is in contrast to the sensitivity of the CD assay which was the assay used to calculate the K_M values previously, and thus we believe the previously reported values to be more accurate.

Reported activation parameters of MR in the aqueous buffers (Tris, Hepes, and phosphate, all at pH 7.5) vary widely (Table 3-5), where even the reported literature values do not correlate well. The activation energy estimated from the slope of the Arrhenius plot (26 ± 1.3 kJ/mol) is within the normal range of 20 – 150 kJ/mol for proteins. The activation enthalpy, determined from the Eyring plot (24 ± 1.2 kJ/mol) is less than the two published values (65 ± 2 kJ/mol; and 133.6 ± 4 kJ/mol). This is at the low end of the expected range for activation enthalpies, 20 – 150 kJ/mol (*10*) and suggests a faster rate. While the non-enzymatic reaction has an activation enthalpy of 130 kJ/mol, we find as expected a significant decrease from this value in the enzymatic reaction. Only one set of data reports a positive activation entropy ($\Delta S^{\neq}$), while the other two data sets report negative activation entropies. A positive activation entropy suggests that the transition state is highly disordered, and the reaction is fast, while a negative activation entropy suggests that the transitions is highly ordered. The free activation energy ($\Delta G^{\neq}$) we calculated is 54 kJ/mol comparable to 56 kJ/mol but almost a third less than 147 kJ/mol.

3.5 Summary and Conclusion

We successfully cloned the mandelate racemase gene, and overexpressed, and purified the his-tagged protein. This his-tagged enzyme has increased activity when compared to published values. While the k_{cat} value is slightly higher when compared to literature values,

the K_M value is significantly different from published values. Therefore, further characterization with a more sensitive polarimeter (lower detection limits and smaller sample volume) is essential. Activation parameters from the temperature data did not correlate well with published values.

3.6 REFERENCES

(1) Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R. S., Kenyon, G. L., and Gerlt, J. A. (1991) Mechanism of the Reaction Catalyzed by Mandelate Racemase .2. Crystal-Structure of Mandelate Racemase at 2.5-a Resolution - Identification of the Active-Site and Possible Catalytic Residues. *Biochemistry 30*, 9264-9273.

(2) Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate Racemase and Muconate Lactonizing Enzyme Are Mechanistically Distinct and Structurally Homologous. *Nature 347*, 692-694.

(3) Mitra, B., Kallarakal, A. T., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase - Importance of Electrophilic Catalysis by Glutamic-Acid-317. *Biochemistry 34*, 2777-2787.

(4) St. Maurice, M., and Bearne, S. L. (2002) Kinetics and Thermodynamics of Mandelate Racemase Catalysis. *Biochemistry 41*, 4048-4058.

(5) Hegeman, G.D., Rosenburg, E.Y., and Kenyon, G.l. (1970) Mandelic Acid Racemase from *Pseudomonas Putida.* Purification and Properties of the Enzyme. *Biochemistry 9*, 4029-4035.

(6) Ransom, S.C., Gerlt, J.A., Powers, V.M., and Kenyon, G.L. (1988) Cloning, DNA Sequence Analysis, and Expression of *Escherichia coli* of the Gene of Mandelate Racemase from *Pseudomonas putida*. *Biochemistry 27*, 540-545.

(7) Bearne, S.L., St. Maurice, M., and Vaughan, M.D. (1999) An Assay for Mandelate Racemase Using High-Performance Liquid Chromatography. *Analytical Biochemistry 269*, 332-336.

(8) Sharp, T.R., Hegeman, G., and Kenyon, G.L. (1979) A Direct Kinetic Assay for Mandelate Racemase Using Circular Dichroic Measurements. *Analytical Biochemistry 94*, 329-334.

(9) Stecher, H., Hermetter, A., and Faber, K. (1998) Mandelate Racemase Assayed by Polarimetry. *Biotechnology Techniques 12*, 257-261.

(10) Bearne, S. L., and Wolfenden, R. (1997) Mandelate Racemase in Pieces: Effective Concentrations of Enzyme Functional Groups in the Transition State. *Biochemistry 36*, 1646-1656.

(11) Whitman, C.P., Hegeman, G., Cleland, W.W., and Kenyon GL. (1985) Symmetry and Asymmetry in Mandelate Racemase Catalysis. *Biochemistry 24*, 3936-3942.

(12) Fee, J.A., Hegeman, G., and Kenyon, G.L. (1974) Mandelate Racemase from Pseudomonas putida. Subunit Composition and Absolute Divalent Metal Ion Requirement. *Biochemistry 13*, 2528-2532.

(13) Fee, J.A., Hegeman G., and Kenyon, G.L. (1974) Mandelate Racemase from Pseudomonas putida. Affinity Labeling of the Enzyme by D,L-α-Phenylglycidate in the Presence of Magnesium Ion. *Biochemistry 13*, 2533-2538.

(14) Burley, R. K. M., and Bearne, S. L. (2005) Inhibition of mandelate racemase by the substrate-intermediate-product analogue 1,1-diphenyl-1-hydroxymethylphosphonate. *Bioorganic & Medicinal Chemistry Letters 15*, 4342-4344.

(15) Schafer, S.L., Barrett, W.C., Kallarakal, A.T., Mitra, B., Kozarich, J.W., Gerlt, J.A., Clifton, J.G., Petsko, G.A., and Kenyon, G.L. (1996) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the D270N Mutant. *Biochemistry 35*, 5662-5669.

(16) Landro, J. A., Gerlt, J. A., Kozarich, J. W., Koo, C. W., Shah, V. J., Kenyon, G. L., Neidhard, D. J., Fujita, S., and Petsko, G. A. (1994) The Role of Lysine 166 in the Mechanism of Mandelate Racemase from Pseudomonas putida: Mechanistic and Crystallographic Evidence for Stereospecific Alkylation by (R)-α-Phenylglycidate. *Biochemistry 33*, 635-643.

(17) Mitra, B., Kallarakal, A. T., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Importance of Electrophilic Catalysis by Glutamic Acid 317. *Biochemistry 34*, 2777-2789.

(18) Kallarakal, A. T., Mitra, B., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the K166R Mutant. *Biochemistry 34*, 2788-2797.

(19) Gerlt, J. A., Kenyon, G. L., Kozarich, J. W., Neidhart, D. J., and Petsko, G. A. (1992) Mandelate Racemase - Understanding the Mechanism of Heterolytic C-H Bond-Cleavage. *Biochemistry 31*, 2188-2188.

(20) Westheimer, F. H. (1995) Coincidences, Decarboxylation, and Electrostatic Effects. *Tetrahedron 51*, 3-20.

(21) Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, Vol. 3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

(22) Murphy, T. (2002) in *Chemical Engineering*, Georgia Institute of Technology, Atlanta.

CHAPTER 4

DIFFERENTIAL SCANNING CALORIMETRY STUDY OF THE IRREVERSIBLE THERMAL DENATURATION OF MANDELATE RACEMASE

4.1 Introduction

Exploring the solar system for extraterrestrial life leads to environments that are different from Earth; such environments tend to be low temperature, non-aqueous environments. We are interested in investigating the events that cause mandelate racemase to become unstable, providing insight into the possibility for enzymatic survival on extraterrestrial planets and moons. The crystal structure of mandelate racemase reveals a tightly packed octamer of identical subunits (*1*). The octameric structure of MR has been described as a tetramer of dimers, where residues of the active site form a 2-fold symmetric subunit (*1*). It has been shown that multimeric proteins can change in oligomericity during the protein denaturation process (*2-4*). With this in mind, we have investigated the thermal denaturation of mandelate racemase to determine the nature of the dissociation as well as to propose a model of

thermal denaturation for this protein. Mandelate racemase has been extensively studied, with available literature data on the enzyme's crystal structure (*5-8*), studies on the reaction mechanism (*5*, *7*, *9-11*), reaction kinetics (*12-14*), and reaction thermodynamics (*10*, *14*, *15*). However, unfolding and possible dissociation of MR have not been investigated. The only member of the enolase superfamily whose denaturation has been studied using differential scanning calorimetry (DSC) is yeast enolase, isozyme 1, a dimeric protein that dissociates by removal of the metal, Mg^{2+} (*16*).

Differential scanning calorimetry (DSC) has been used to provide insight into the energetics, kinetics, and the mechanism of reversible unfolding of enzymes using equilibrium thermodynamics as well as the development of thermal denaturation models for enzymes that unfold irreversibly. Proteins that undergo fully reversible two-state folding/unfolding transitions tend to be relatively small with molecular weights less than 20 kDa (*2*, *17*). Larger multimeric proteins, however, have a propensity for more complicated unfolding transitions due to the independent unfolding and dissociation of separate subunits (*3*, *18*, *19*). Consequently, thermal denaturation of larger proteins in many cases leads to an irreversible DSC thermogram, in which case analysis using equilibrium thermodynamics alone is not applicable (*3*, *20-25*). As shown

below, such is the case with MR, so determining a model for the irreversible thermal denaturation is the goal of this work.

4.2 Background

4.2.1 DSC to measure enzyme thermostability

Protein stability can be affected by thermodynamic, kinetic and thermal aspects (*26*). Thermodynamic stability describes protein unfolding on the global level (*27*). The free energy difference between the native (N) state and unfolded, or nonstructured (U) state is described by the equilibrium property ΔG, with the simplest model of unfolding represented as follows:

$$N \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} U$$

where k_1 is the unfolding rate constant and k_{-1} is the folding rate constant. There are most likely intermediate structures present between these two states, especially in the case of multimeric proteins. The change in free energy of unfolding describes the amounts of the native and unfolded states present at each temperature, according to equation 4-1; however, this measurement is dependent on the reversibility of the reaction.

$$\Delta G_{unfolding} = -RT \ln K \qquad 4\text{-}1$$

where R is the universal gas constant, T is temperature, and K is the equilibrium constant. The major driving force for protein unfolding is from

hydrophobic effects, which in turn causes ΔG to be curved function of temperature (*28*). Both enthalpy and entropy become a function of temperature, this change can then be described by the change in heat capacity as shown in equation 4-2:

$$\Delta C_p = \left(\frac{\partial \Delta H(T)}{\partial T} \right)_p = \left(\frac{T \partial \Delta S}{\partial T} \right)_p . \qquad 4\text{-}2$$

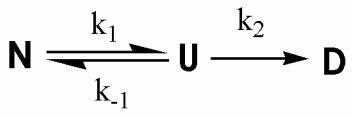
Using these relationships, the free enthalpy of unfolding, ΔG, can be determined by plotting ΔG versus temperature, the so-called protein stability curve. The temperature where ΔG is equal to 0, i.e. the point where the native and unfolded states are equally populated, is the melting temperature T_M. This maximum temperature is a measure of thermal stability.

Differential scanning calorimetry (DSC) is an experimental technique that can be used to characterize temperature-induced conformational changes in the protein (*29*). DSC measures the differences of heat uptake between a sample and reference cell. The resulting thermogram shows the excess heat capacity (C_p^{ex}), which is determined by subtracting the heat capacity of the reference cell from that of the sample cell, as a function of temperature. The calorimetric enthalpy (ΔH_{cal}) is then computed by integrating the area under the curve. A typical scan involves heating with subsequent cooling of the protein. Reversibility of the process is determined by including a second

scan after the protein is cooled. If there is a identical thermogram the denaturation process is calorimetrically reversible. Many small globular proteins, such as lysozyme, result in calorimetrically reversible processes, in which case thermodynamic stability can then be explored. However, if the system is not calorimetrically reversible, this information cannot be extracted.

4.2.2 Analysis of irreversible thermal denaturation

Irreversible protein denaturation has not been as well studied as the reversible case (*3*). Previously, irreversible denaturation has been shown to occur according to the Lumry-Eyring model:



where k_2 is the rate constant of the irreversible denaturation. The U to D step is the first order chemical degradation, while U to N is the protein folding back to the native state. The denatured (D) state is composed of a mixture of conformations, where the presence of some partially compact structure and secondary elements may still be present (*30*). Until about the mid 1990's many researchers regarded the irreversible step as insignificant during the DSC transition (*18*, *23*, *31*, *32*). However, since this time it has been shown that a scan-rate dependence of the DSC transition is evidence for kinetic control and such a transition cannot be characterized by thermodynamic equilibrium analysis (*33-35*).

Several theoretical models have been proposed for the analysis of irreversible denaturation (*2-4*, *20*) for both monomeric and multimeric proteins. In this thesis we explore the multimeric protein MR, so it is such models that we will further investigate. Prior to development of the models it is necessary to discuss the possible situations that can arise in the DSC transitions; for simplicity we will discuss this in terms of the Lumry-Eyring model.

Each situation is based on T*, the characteristic temperature for the irreversible process, i.e. the temperature where the kinetic constant, k_2, is equal to unity and T_M is the temperature where the equilibrium constant is equal to unity.

- The first possible situation involves the case where $T^* >> T_M$. In this situation although the protein is irreversibly unfolding by definition standards, the protein is reversibly unfolding at a temperature lower than the irreversible step, resulting in a calorimetric trace similar to that of reversible unfolding; in this case, application of equilibrium thermodynamics is valid.
- The second case is when $T^* \approx T_M$, at which point the irreversible step is fast but not dominating in the temperature range of the reversible unfolding. At this point, the DSC transition is significantly affected by the irreversible unfolding. In turn,

applying equilibrium thermodynamics will lead to grave errors (*33*).

- In the third case when $T^* << T_M$ the irreversible step occurs at a temperature much lower than the melting temperature. Here, the irreversible denaturation will cause the apparent T_M from DSC to be greater than the T_M for the unfolding process, hence, application of equilibrium thermodynamics would lead to gross errors (*33*).

4.2.3 Models

Through scan rate dependence and irreversibility of the MR thermal transition, it was determined that thermodynamic parameters cannot be obtained; therefore only kinetic parameters can be acquired. Many theoretical models have been previously described for kinetic analysis of thermal transitions (*2-4*, *20*, *36-38*). We began by fitting the excess heat capacity data to the simple two-state irreversible model and then further investigated the consecutive two-step irreversible model as well as the Lumry-Eyring model.

4.2.3.1 Two-State Irreversible Model

The simplest model is the irreversible denaturation of native state protein, with one-step dissociation as shown in scheme I:

$$N_n \xrightarrow{k} nD\,, \qquad \text{Scheme 1}$$

where N and D are the native and denatured states of the protein, k is the first order rate constant, and n is the oligomericity of the native protein. Irreversible DSC profiles have been analyzed in terms of this model by Freire et. al. (*35*), Sanchez-Ruiz (*3, 29*), Sanchez-Ruiz et. al. (*33*), and in more detail by Kurganov et. al. (*20*), who discussed the criteria of validity for this model. As this is a limiting case of the Lumry-Eyring model, the rate constant ***k*** is assumed to follow the Arrhenius law, equation 4-3:

$$k = \exp\left(\left(\frac{E_a}{R}\right)\left(\frac{1}{T*} - \frac{1}{T}\right)\right) \qquad 4\text{-}3$$

where E_a is the apparent energy of activation, T is the absolute temperature, T^* is the temperature at which k = 1 min^{-1}. The rate equation for this model is equation 4-4

$$\frac{d[N]}{dt} = -k[N] \qquad 4\text{-}4$$

where [N] is the concentration of protein in the native state. If the temperature is a variable parameter and the rate of variation of the temperature is constant, then equation 4-4 becomes equation 4-5

$$\frac{d[N]}{dT} = -\left(\frac{1}{\nu}\right)k[N] \qquad 4\text{-}5$$

where ν is the scan rate. Integrating equation 4-5 yields an expression for the relative amount of the N state as shown in equation 4-6:

$$\frac{[N]}{[N]_0} = \exp\left\{-\frac{1}{\nu}\int_{T_0}^{T}\exp\left[\frac{E_a}{R}\left(\frac{1}{T*} - \frac{1}{T}\right)\right]dT\right\} \qquad 4\text{-}6$$

where $[N]_0$ is the initial concentration of N at the initial scanning temperature, T_0. Taking the native state as the reference state, the excess enthalpy $\langle \Delta H \rangle$ can be defined as equation 4-7:

$$\langle \Delta H \rangle = \left(1 - \frac{[N]}{[N]_0}\right) \Delta H \qquad 4\text{-}7$$

where ΔH is the denaturation enthalpy. Again taking the native state as the reference state, the excess heat capacity yields equation 4-8:

$$C_p^{ex} = \frac{1}{v}\frac{\Delta H}{n} \exp\left\{\frac{E_a}{R}\left(\frac{1}{T^*} - \frac{1}{T}\right)\right\} \exp\left\{-\frac{1}{v}\int_{T_0}^{T} \exp\left[\frac{E_a}{R}\left(\frac{1}{T^*} - \frac{1}{T}\right)\right] dT\right\} \qquad 4\text{-}8$$

The C_p^{ex} values can then be plotted against the scanning temperature range to result in an asymmetric curve with a maximum temperature. This model curve can then be compared to experimental data to determine if the model fits the data.

The monomolecular form of this model has been extensively discussed in literature (*3, 4, 29, 35, 36, 39*) and Sanchez-Ruiz et. al. developed four methods to estimate the apparent activation energy E_a (*33*). Agreement between the different values calculated for E_a is used as a convenient test for the validity of the two-state irreversible model and this has been used extensively in literature to determine the validity of this model, i.e. if this model is applicable (*21, 22, 33, 34, 40, 41*). It should be noted that some of these criteria contain approximations and some have

limited accuracy, but the combination of all of the criteria yields a viable method for determining the validity of this model. All of these criteria are derived from the above set of equations (*33*).

The first criterion calculates the activation energy of the irreversible transition from an Arrhenius plot of the first-order kinetic constant determined at each temperature by:

$$k = \frac{vC_p^{ex}(T)}{\Delta H_{cal} - \Delta H(T)} \qquad 4\text{-}9$$

where C^{ex}_p is the excess heat capacity, ΔH_{cal} is the total enthalpy change of the process, and ΔH is the enthalpy change at a given temperature. The activation energy can then be determined from the Arrhenius slope, which is equivalent to the expression ($-E_a/R$).

The second criterion uses the temperature dependence on the heat evolution as determined by the following:

$$\ln\left(\ln\left[\frac{\Delta H_{cal}}{\Delta H_{cal} - \Delta H(T)}\right]\right) = \frac{E_a}{R}\left(\frac{1}{T_{m,app}} - \frac{1}{T}\right). \qquad 4\text{-}10$$

Plotting $\ln[\ln(\Delta H_{cal}/(\Delta H_{cal}-\Delta H(T))]$ versus 1/T is expected to result in a straight line with a slope of $-E_a/R$ for each scanning rate.

The third criterion based on the kinetic model predicts the temperature value corresponding to the maximum temperature of the heat capacity curve, T_m, which changes with the scan rate according to equation 4-11:

$$\ln\left(\frac{\nu}{T_m^2}\right) = \ln\left(\frac{AR}{E_a}\right) - \left(\frac{E_a}{R}\right)\left(\frac{1}{T_m}\right) \qquad 4\text{-}11$$

A plot of ln (ν /$T_{m,app}^2$) vs. 1/$T_{m,app}^2$ should give a linear dependency with the slope equal to E_a/R. As the previous methods are dependent on a range of temperatures resulting in an E_a for each scan rate, this criterion involves only one data point for each scan rate leading to less variation in the activation energy.

Through relation of the Arrhenius equation to the rate equation, at $T = T_m$, the apparent activation energy, E_a, can be calculated using equation 4-12, where C_p^m is the maximum heat capacity of the calorimetric trace:

$$E_a = \frac{eRC_p^m T_m^2}{\Delta H_{cal}} \qquad 4\text{-}12$$

4.2.3.2 Lumry-Eyring Model

The simple one-step irreversible model has been considered to be only an approximation (*20*), because proteins can likely have more complicated denaturation pathways with one or several intermediate states. Thus, more detailed models might better describe the denaturation of mandelate racemase. The Lumry-Eyring model is a frequently used model that describes denaturation as reversible unfolding followed by an irreversible denaturation event. It has been suggested that the Lumry-Eyring model is more realistic when making the assumption

of a rapid equilibrating first step with concurrent dissociation (*33*). In this model the multimeric native protein, *N*, undergoes a two-state reversible unfolding with simultaneous dissociation into n monomers with the unfolded species *U* undergoing an irreversible alteration to yield a final state, *D*.

$$N_n \overset{K}{\longleftrightarrow} nU \xrightarrow{k} D. \qquad \text{Scheme II}$$

The rapid equilibrium assumption implies there is a very small amount of U present at all times, represented by:

$$K = \frac{[U]^n}{[N_n]} = nM^{(1-n)} \frac{C_U^n}{C_N} \qquad 4\text{-}13$$

where M is the molecular weight of the protein per monomer basis. It is further assumed that dissociation and unfolding occur simultaneously, and the irreversible step follows first-order kinetics that change with temperature according to the Arrhenius law, as follows:

$$k_a = \exp\left(-\frac{E_a}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right) \qquad 4\text{-}14$$

where k_a is the apparent rate constant. Another assumption is that the heat capacity does not change between the unfolded and native state. The rate equation for the final step of this system, at a constant scanning rate, can be defined as follows:

$$\frac{d[D]}{dt} = k[U] \qquad 4\text{-}15$$

Combining equation (4-15) and (4-13) will then yield:

$$\frac{d[D]}{dt} = k_a[N]^{1/n} \qquad 4\text{-}16$$

where k_a is an apparent rate constant defined as:

$$k_a = k\left(\frac{K}{n}\right)^{1/n} M^{((n-1)/n)} \qquad 4\text{-}17$$

Since k_a is proportional to $kK^{1/n,}$ and both k and K are assumed to show exponential dependence with 1/T, the temperature dependence of k_a can be expressed by the Arrhenius equation:

$$k_a = \exp\left(-\frac{E_a + \Delta H_U / n}{R}\left(\frac{1}{T} - \frac{1}{T^*}\right)\right) \qquad 4\text{-}18$$

From here the apparent energy of activation E_a can be defined as:

$$E_a = E + (\Delta H_U / n) \qquad 4\text{-}19$$

Incorporating equation 4-19 into 4-18:

$$k_a = \exp\left(\frac{E_a}{RT^*}\right)\exp\left(-\frac{E_a}{RT}\right) \qquad 4\text{-}20$$

Hence, equation 4-15 can be related to the native protein concentration:

$$\frac{d[N]}{dt} = -\frac{d[D]}{dt} = -k_{app}[N]^{1/n} \qquad 4\text{-}21$$

In the DSC experiment the temperature changes with time according to a constant scanning rate ν = dT/dt, incorporating this into equation 4-21 reveals:

$$\frac{d[N]}{dt} = -\left(\frac{1}{\nu}\right) k_a(T)[N]^{1/n} \qquad 4\text{-}22$$

To solve equation 4-22 we then integrate numerically using a fourth-order Runge-Kutta algorithm from a low temperature T_0, where the rate is negligible and [N] = C_T, to a temperature T. The excess heat capacity can then be obtained:

$$C_p^{ex} = -\frac{\Delta H}{[N]+[D]}\frac{d[N]}{dT} \qquad 4\text{-}23$$

The denaturation enthalpy, ΔH, is assumed constant

4.2.3.3 Consecutive two-step irreversible model

By assuming that the rate of the second step of the Lumry-Eyring model is much greater than the reverse reaction rate of the first step, a model using two irreversible transitions results (*42*).

$$N_n \xrightarrow{k_1} nU \xrightarrow{k_2} D \qquad \text{Scheme III}$$

Taking into account the scan rate where the rate of the variation of temperature is constant (ν = dT/dt), the rate equations for this system can be developed as follows:

$$\frac{d[N]}{dT} = -\frac{k_1}{\nu}[N], \qquad 4\text{-}24$$

$$\frac{d[U]}{dT} = \frac{1}{\nu}\left[nk_1[N] - k_2[U]\right], \qquad 4\text{-}25$$

$$\frac{d[D]}{dT} = \frac{k_2}{\nu}[U], \qquad 4\text{-}26$$

where [N], [U], and [D] are the molar concentrations of the native, unfolded and denatured states, respectively, k_1 is the rate constant for the first step and k_2 is the rate constant for the second step. With the assumption that the rate constants follow the Arrhenius law, the rate constants for both steps can be defined in the following manner:

$$k_1 = \exp\left(-\frac{E_1}{R}\left(\frac{1}{T_1^*}-\frac{1}{T}\right)\right) \quad , \qquad 4\text{-}27$$

$$k_2 = \exp\left(-\frac{E_2}{R}\left(\frac{1}{T_2^*}-\frac{1}{T}\right)\right) \quad , \qquad 4\text{-}28$$

where T_1^* and T_2^* are the characteristic reference temperatures where the rate constants k_1 and k_2 are equal to 1 min^{-1} for the first and second transitions, respectively. Integrating equation (8), the amount of protein in the N state can be expressed as a function of temperature:

$$[N] = [N_0]\exp\left(-\frac{1}{\nu}\int_{T_0}^{T} k_1 dT\right), \qquad 4\text{-}29$$

where $[N_0]$ is the initial concentration of the protein in the native state. Substituting equation (13) into equation (8):

$$\frac{d[U]}{dT} + \frac{k_2}{\nu}[U] = \frac{nk_1[N_0]}{\nu}\exp\left(-\frac{1}{\nu}\int_{T_0}^{T} k_1 dT\right). \qquad 4\text{-}30$$

Assuming that $N_0 = 0$ at $T = T_0$:

$$[U](T) = \frac{n[N_0]}{\nu}\exp\left(-\frac{1}{\nu}\int_{T_0}^{T} k_2 dT\right)\int_{T_0}^{T}\left[k_1\exp\left(\frac{1}{\nu}\int_{T_0}^{T}(k_2 - k_1)dT\right)\right]dT. \quad 4\text{-}31$$

The excess enthalpy is then determined by the sum of the excess enthalpy change for both steps, expressed as:

$$C_p^{ex} = \frac{d\langle \Delta H \rangle}{dT} = \frac{\partial H}{\partial U}\frac{\partial U}{dT} + \frac{\partial H}{\partial F}\frac{\partial F}{dT}. \qquad 4\text{-}32$$

The C_p^{ex} can then be defined as:

$$C_p^{ex} = \frac{\Delta H_1 n k_1}{v}[N] + \frac{\Delta H_2 k_2}{v}[U]. \qquad 4\text{-}33$$

4.2.4 Thermostability of mandelate racemase

The thermostability of mandelate racemase has never been studied, however, a member of the enolase superfamily has been. A DSC study of yeast enolase I was conducted, taking into account effects of the metal ions, substrate/product, substrate analogs, and chaotropic anions (*16*). This enzyme, however, is very different from mandelate racemase in many ways. The amino acid sequences between MR and enolaseI shows 20% identity. Also, enolase I is a dimer while mandelate racemase is an octamer. As both belong to the same superfamily, the two proteins have similar tertiary structures, although not superimposable as MR is with MLE. Also, both contain metals but interestingly, the metal cofactor in yeast enolase 1 is located between the two subunits of the protein. Therefore, a monomeric form of this protein can be prepared by extracting the metal, providing both the multimeric and monomeric forms for DSC analysis. It was concluded that the enzyme is more stable in the

dimeric form and enzyme stabilization is increased with increased strength of the metal binding (*16*).

However, the metal in mandelate racemase does not link the identical subunits. Since the crystal structure has been solved some knowledge exists about how the subunits interact. It has been determined that interfaces between two-fold related subunits are more intimate than between four-fold interfaces, so the structure can be described as a tetramer of dimers (*43*) (Figure 4-1).

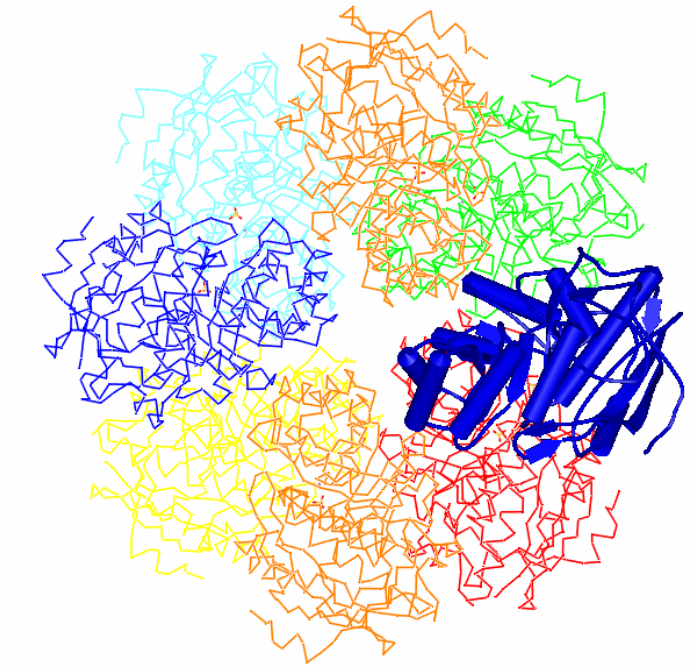


Figure 4-1. Structure of mandelate racemase showing interaction of subunits.

4.3 Materials and Methods

4.3.1 Materials

Tris(hydroxymethyl)-aminomethane was obtained from Sigma-Aldrich (St. Louis, Missouri), and $MgCl_2$ was from Fisher Chemical (Fair Lawn, New Jersey).

4.3.2 DSC measurements

DSC was performed using an NDSC-II instrument (Calorimetry Sciences, Lindon, Utah). The scans were run at 2, 1.5, 0.75, 0.5, and 0.25 °C/min from 30 °C to 90 °C at a constant pressure of 3 atm and then analyzed using the CpCalc software available with the instrument. Protein concentrations measured include 1, 2, 4, and 6 mg/mL, as determined using a Bradford assay.

4.3.3 CD measurements

The ellipticity at 280 nm was measured over a change in temperature using a J-815 Circular Dichroism Spectropolarimeter (Jasco, Easton, Maryland). Absorbance was measured every 5°C in the temperature range of 30°C to 75°C. A wavelength of 222 nm was employed and measurements were taken in the temperature range from 30°C to 90°C.

4.3.4 Analysis of data

The chemical background (buffer) scan was subtracted from the MR DSC scan to obtain the net heat rate from the raw data files. The net

heat rate was then divided by the scan rate to determine the heat capacity for each point, and plotting the heat capacity versus the temperature yielded a revised thermogram. The excess molar heat capacity was determined by subtracting the pre-thermal transition baseline until pre- and post-baselines resulted in the same value, after which we switched to subtracting the post-thermal transition baseline. The maximum temperature ($T_{m,app}$) was determined as the corresponding temperature of the largest excess heat capacity value, or the temperature at the peak of the DSC thermogram. The experimental calorimetric enthalpy was calculated as the total integrated area under the thermogram peak after the baseline corrections. DSC data at each condition of MR concentration or scan rate was taken in triplicate.

The models were fit to the experimental excess molar heat capacity versus temperature data using computer algorithms implemented in MATLAB that minimize the least squares error between the experimental and model data. The accuracy of fitting was determined by the correlation coefficient, r, as in Equation 4-34:

$$r = \sqrt{\frac{1 - \sum_{i=1}^{n} (y_i - y_i^{calc})^2}{\sum_{i=1}^{n} (y_i - y_i^{m})^2}} \qquad 4\text{-}34$$

The models in the form of dy/dt = f(t) were solved with an adaptive Gauss/Lobatto quadrature method (*44*). The ordinary differential

equations (ODE) were solved using the Runge-Kutta (4,5) algorithm using Dormand-Prince pairs (*45*). The simplex search and sequential quadratic programming methods were the algorithms used to minimize the objective function.

4.4Results and Discussion

4.4.1 Dependence of maximum temperature on scan rate

The thermal unfolding for MR after the subtraction of the chemical background scan is shown in Figure 4-2 at a scanning rate of 1 K/min and 1 mg/mL. In all cases, a second scan did not result in a transition, and aggregation was visually observed by precipitation when samples were removed from the instrument. The excess heat capacity versus temperature profile contains a single transition with a maximum temperature, $T_{m,app}$, of 66 °C for this scanning rate and concentration. No DSC transition is observed upon the second scan in Figure 4-2, which allows the conclusion that the MR thermal denaturation process is irreversible. Through further experimentation, we identified aggregation at temperatures near the start of the transition, a second sign of an irreversible process. Since the process is irreversible, the thermodynamic parameters cannot be obtained for the thermal denaturation of this enzyme.

Thermograms of 1 mg/mL MR were also obtained at scanning rates of 0.25, 0.5, 0.75, 1.0, 1.5, and 2.0 K/min, the maximum scan rate

attainable with the DSC instrument. The thermograms of MR at varying scan rates, shown in Figure 4-3, reveal an observable shift in the maximum transition temperature with increasing scan rate. As the irreversible thermal transition of mandelate racemase is dependent on the scan rate with the maximum temperature increasing with an increase in scan rate, the process is kinetically controlled (*2*, *3*). The scan limit, which is a scan rate where the thermal transition will reach a plateau, is not observed, therefore, it is higher than the maximum attainable scan rate (2 K/min) of the instrument (*37*, *38*).

4.4.2 Protein concentration dependence

Mandelate racemase is composed of eight identical subunits, therefore, thermal denaturation may also involve dissociation of these subunits, which can occur in a variety of scenarios (schemes IV – VI).

$$N_8 \overset{K}{\longleftrightarrow} 8U \xrightarrow{k} 8D, \qquad \text{Scheme IV}$$

$$N_8 \overset{K}{\longleftrightarrow} 4U_2 \xrightarrow{k} 4D_2, \qquad \text{Scheme V}$$

$$N_8 \overset{K}{\longleftrightarrow} 2U_4 \xrightarrow{k} 2D_4. \qquad \text{Scheme VI}$$

It has been shown that multimeric proteins undergoing two-state reversible unfolding with simultaneous dissociation into monomers exhibit an increase in the maximum temperature with an increase in total protein concentration (*46*). This phenomenon has also been observed when modeling irreversible thermal transitions (*3*, *23*). The occurrence of

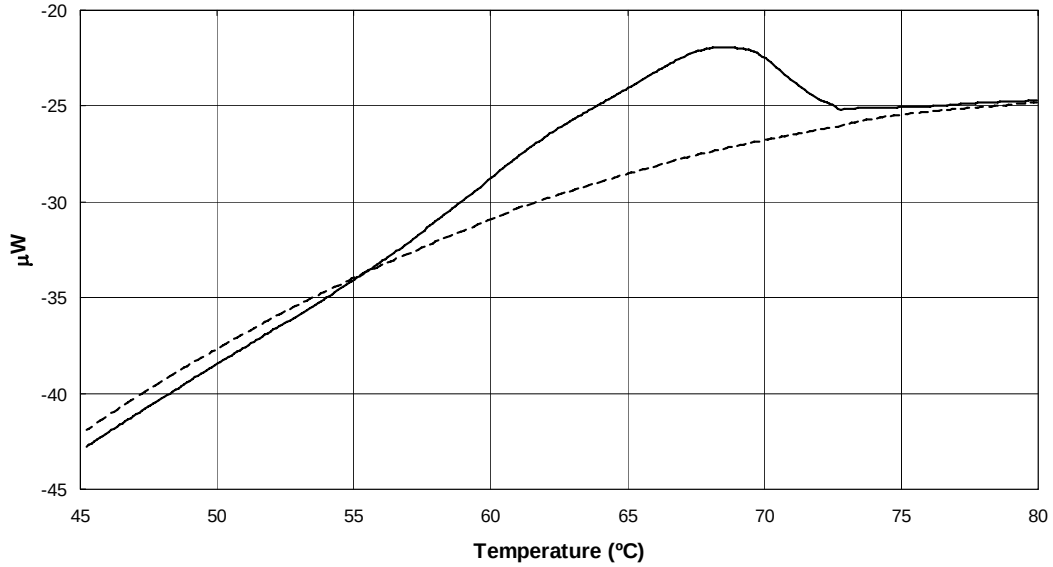


Figure 4-2. Thermal transition of mandelate racemase (scan rate: 1 K/min, = 1 mg/mL), with subsequent rescan.

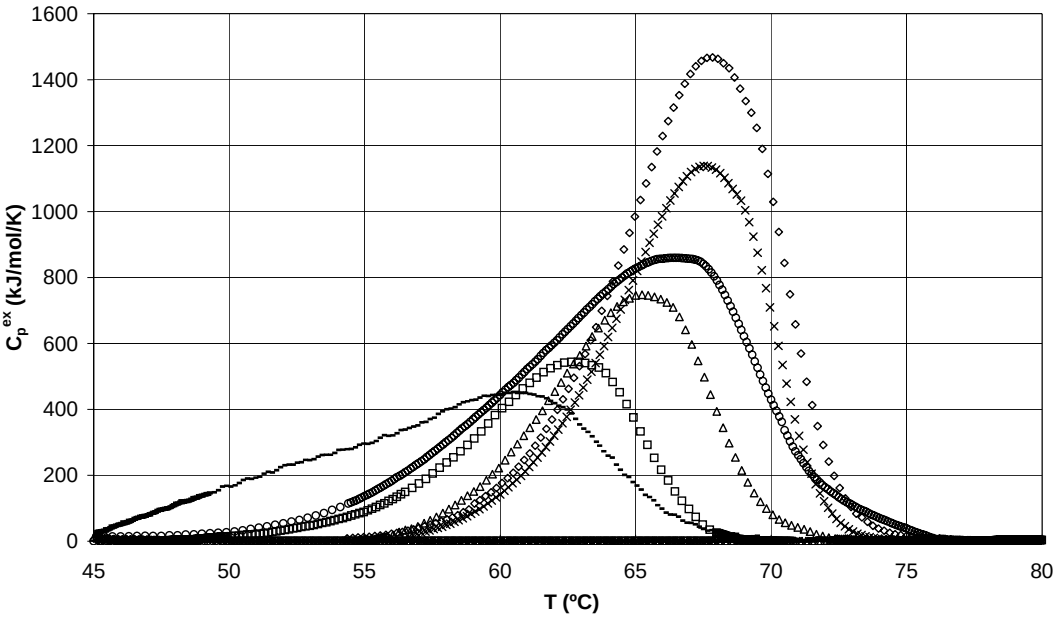


Figure 4-3. DSC thermograms for mandelate racemase (1 mg/mL) at varying scan rates. (◊) 2 K/min; (✕) 1.5 K/min; (○) 1 K/min; (Δ) 0.75 K/min; (□) 0.5 K/min; (—) 0.25 K/min.

dissociation can be determined by investigating the effect of protein concentration on the thermogram (*2*, *3*).

No concentration dependence was observed at a scan rate of 0.25 °C/min. The maximum temperature for all concentrations ranged between 60.3°C and 60.6°C, which was also reproduced using CD at a scan rate of 0.25 °C/min and concentration of 0.05 mg/mL. We were unable to measure thermal transitions using CD for any concentrations above 0.05 mg/mL, which was also the lowest concentration still measurable within instrumental error.

4.4.3 Investigation of enthalpy values

The van't Hoff enthalpy, ΔH_{vH}, can be obtained from the DSC trace using the following relationship (*23*):

$$\Delta H_{vH} = \frac{ART_m^2 C_p^{ex}}{\Delta H} \qquad 4\text{-}36$$

where A is a constant dependent on n, this value was taken from Manly et. al (*23*). In the case of the two-state reversible unfolding model, the calorimetric enthalpy (ΔH) is equal to the van't Hoff enthalpy (ΔH_{vH}). Therefore, in the case of two-state irreversible unfolding the expectation is that the calculated ΔH_{vH} does not equal the calorimetric enthalpy. Reported in Table 4-1 are the values of the calculated van't Hoff enthalpies, ΔH_{vH}, in comparison with the measured enthalpies, ΔH. In all cases the van't Hoff enthalpy is significantly less than the experimental

enthalpies. As further proof that thermodynamic quantities cannot be extracted from the MR thermal transitions, calculated van't Hoff enthalpies, ΔH_{vH}, were compared to measured enthalpies, ΔH. If these values are statistical the same, it is reasonable to extract thermodynamic data, even though irreversibility is observed (*29*). In all cases, there is a drastic difference between these values, therefore further evidence that the thermal transition of MR is kinetically controlled. This shows that the irreversible case takes more energy to denature, than if it was the reversible case.

Table 4-1. Comparison of calculated van't Hoff enthalpies and measured calorimetric enthalpies.

ν(K/ min)	[MR] (mg/mL)	ΔH (kJ/mol)	Cp_m^{ex} (kJ/mol/K)	T_m (K)	ΔH_{VH} (kJ/mol)
0.25	1	5405 ± 1189	464 ± 84	333.8 ± 0.67	318 ± 98
0.5	1	2770 ± 609	356 ± 64	337.8 ± 0.92	488 ± 103
0.75	1	4098 ± 902	568 ± 93	338.3 ± 1.1	515 ± 123
1	1	1293 ± 284	1685 ± 305	340.7± 1.5	540 ± 83
1.5	1	8132 ± 1789	1138 ± 423	338 ± 0.99	267 ± 56
2	1	10284 ± 2262	1395 ± 256	339.1 ± 0.48	519 ± 133

4.4.4 Application of the three kinetic models

The two-state irreversible model is the simplest model to describe irreversible thermal denaturation. A profile of excess heat capacity versus temperature was created using a MATLAB program we developed to fit the model to the experimentally obtained MR DSC thermograms. As we

were interested in changes in multimericity of MR as it undergoes thermal denaturation, we explored the expected shape of the curves with fixed multimericity values using the two-state irreversible model, shown in Figure 4-4. It has been observed that the DSC thermograms in cases of multimeric proteins tend to be asymmetric, with the asymmetry increasing with increasing multimericity (*2*, *3*). This same phenomenon is observed in Figure 4-4, where higher multimericities display a sharp drop in the right side of the peak. The experimentally measured plots of excess heat capacity versus temperature and calculated plots using a multimericity of unity (n=1) superimpose far better than plots calculated with higher multimericities.

The r-values for fitting the two-state irreversible model (Scheme 1) to the MR DSC data at varying scan rates and protein concentrations are found in Table 4-2, with the graphical fit of this model to the MR DSC data shown in Figure 4-5 and 4-6. For these fits, the multimericity was not fixed but calculated for the best fit. The overall multimericity value n was calculated to be 0.91 $\pm$ 0.09 and did not change with scan rate providing further evidence that no dissociation seems to occur during irreversible denaturation.

As previously noted, four different criteria have been previously developed to determine if the two-state model validly applies to a given set of denaturation data (*33*). The first criterion is a check for agreement

of the rate constants over the different scan rates. The rate constant was evaluated for all data in the interval $(T_{m,app} - 5) \leq T \leq (T_{m,app} + 5)$ according to Equation 4-7, and the Arrhenius plots can be found in Figure 4-7A.

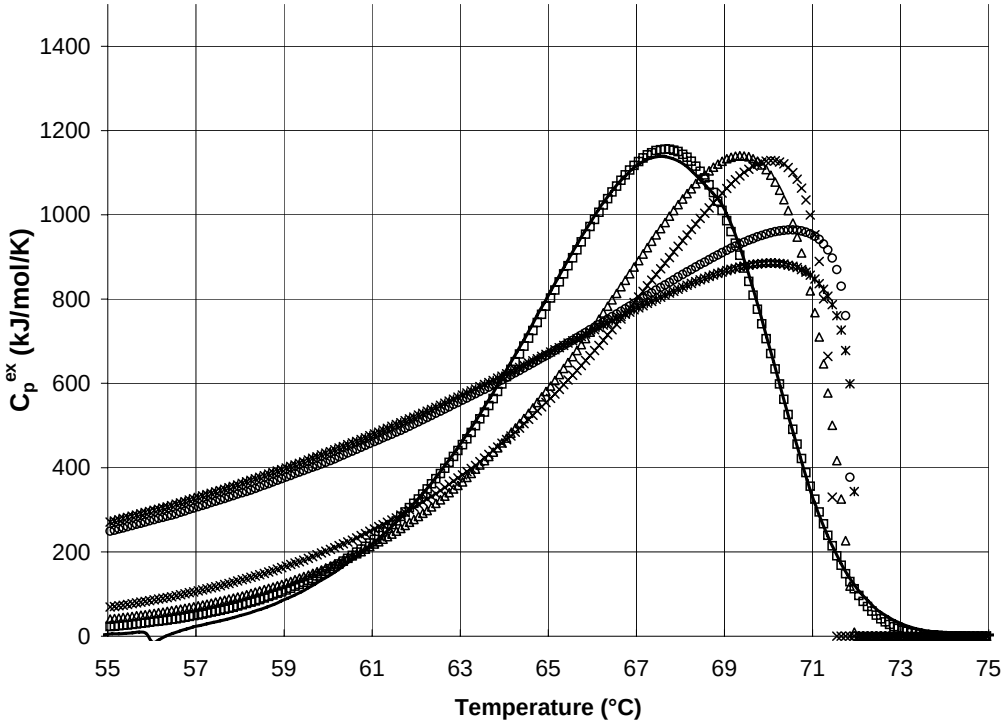


Figure 4-4. Comparison of curves expected for various multimercities using the two-state model. (◊) 1.5 K/min, 1 mg/mL; (?) n=1; (?) n=2; (✕) n=4; (∗) n=6; and (?) n=8.

An activation energy was determined from the slope $-E_a/R$ for each scan rate; the mean and standard deviation is 347.3 ± 25.9 kJ/mol. The second criterion explores the dependence of the heat evolved on the temperature using Equation 4-10; the calculated mean activation energy with standard deviation for this criterion is 317.0 ± 20.5 kJ/mol (Figure 4-7B).

Table 4-2. The r-values for fitting the three irreversible models to the MR DSC data.

		Experimental			Scheme I				Scheme II	Scheme III
ν	Mg/mL	T_m^{app} (°C)	ΔH_{ex} (kJ/mol)	ΔH (kJ/mol)	E_a (kJ/mol)	T* (C)	n	r	R	r
0.25	1	60.8 ± 0.9	5392 ± 253	6220 ± 356	192 ± 9.5	85 ± 1	0.9 ± 0.04	9.3 ± 0.4	25 ± 4	25 ± 4
0.5	1	62.6 ± 1.2	2747 ± 99	2781 ± 219	332 ± 14	83 ± 2	1 ± 0.7	2.4 ± 0.3	5.2 ± 2	5.3 ± 2
0.75	1	63.7 ± 1.4	4130 ± 161	4196 ± 185	365 ± 32	78 ± 2	0.8 ± 0.06	6.9 ± 0.9	14.3 ± 3.2	12.2 ± 1.3
1	1	64.5 ± 1.1	7711 ± 1970	8286 ± 2108	302 ± 40	87 ± 1	0.9 ± 0.06	15.4 ± 2.3	22.9 ± 5	24.6 ± 18
1	2	65.3 ± 0.9	7686 ± 812	7987 ± 1213	321 ± 38	80 ± 3	0.8 ± 0.08	1.1 ± 0.2	8.1 ± 2	25.6 ± 13
1	4	67.3 ± 1.2	7006 ± 832	7655 ± 987	329 ± 17	81 ± 3	0.9 ± 0.08	12 ± 0.6	16.3 ± 5	26.9 ± 12
1	6	66.9 ± 1.7	5758 ± 306	6469 ± 456	353 ± 13	78 ± 2	0.8 ± 0.01	4.2 ± 1.6	27.9 ± 5	21.8 ± 16
1.5	1	65.2 ± 0.7	8158 ± 152	8179 ± 159	369 ± 2.5	79 ± 0.6	1 ± 0.08	13.7 ± 0.5	118 ± 22	17.5 ± 3.4
2	1	66.4 ± 1.4	10158 ± 1152	10931 ± 1779	363 ± 19	79 ± 4	0.9 ± 0.06	15.2 ± 1.1	224 ± 28	27.6 ± 3.1

The third criterion determines the relationship between maximum temperature and scan rate from Equation 4-11 (one data point for each scan rate). The activation energy obtained from the slope of the plot is 388.7 ± 58.2 kJ/mol (Figure 4-7C). The fourth criterion calculates the activation energy using the heat capacity at the maximum temperature using Equation 4-12. The average activation energy for each scan rate is 349.4 ± 30.0 kJ/mol.

To evaluate the agreement between the four criteria, we used the statistical method analysis of variance between groups (*34*). Using the ANOVA method we found the average activation energy values calculated by each of the four criteria (Table 4-3) to be statistically the same, with the overall average activation energy of 350 ± 30 kJ/mol. The good agreement between the four criteria leads us to conclude that use of this model is valid for the experimental data.

Although the experimental data can be described by this model, it has been suggested that the real mechanism is most likely much more complex (*20*, *39*). Both of the models described in scheme II and scheme III were also fitted to the MR thermograms using the computational methods as described. The r-values for both of these can be found in Table 4-2, with the graphical representation shown in Figures 4-7 and 4-8. Neither the Lumry-Eyring model (scheme II) nor the consecutive two-step irreversible model (scheme III) fits the experimental data as well as the

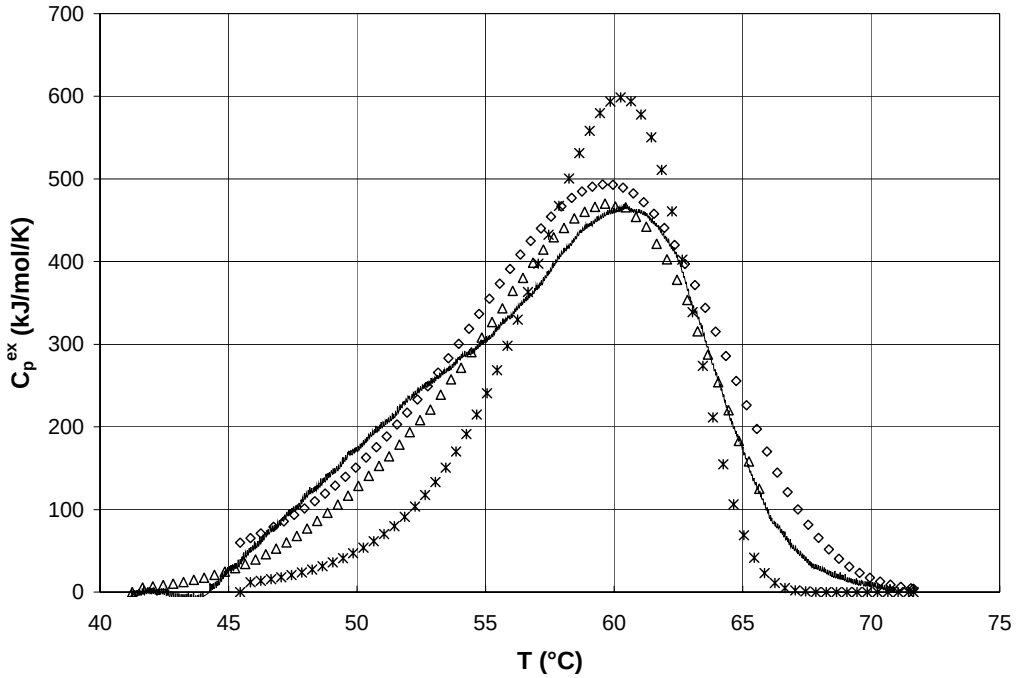


Figure 4-5. Comparison of theoretical and experimental C_p^{ex} vs. T curves for scan rate of 0.25 K/min and 1 mg/mL MR. (Δ) two-state irreversible model, (∗) Lumry and Eyring model, (◊) two irreversible steps model.

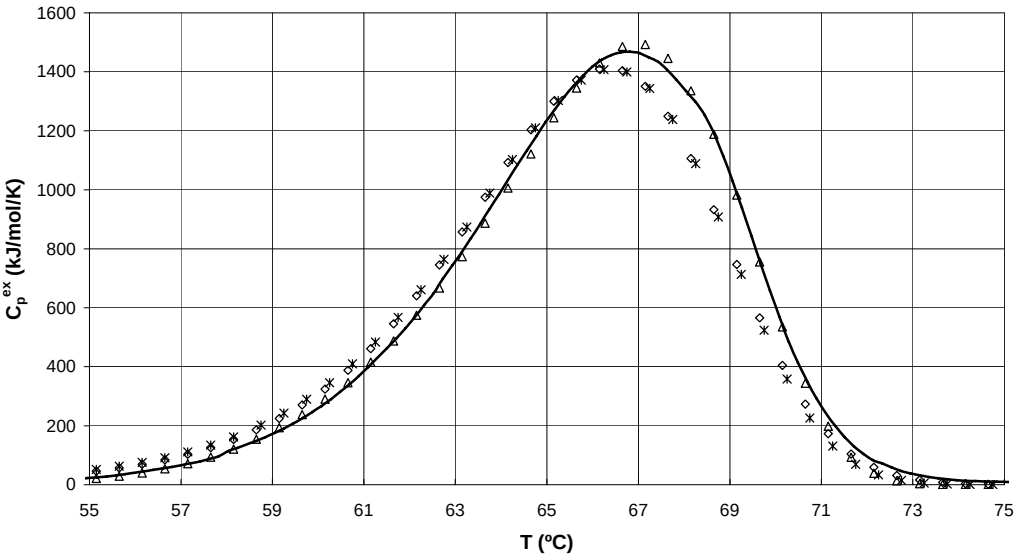


Figure 4-6. Comparison of theoretical and experimental excess heat capacity versus temperature curves for scan rate of 2 K/min and 1 mg/mL MR. (Δ) two-state irreversible model, (∗) Lumry and Eyring model, (◊) two irreversible steps model.

two-state irreversible model. It is observed in the Lumry-Eyring model that the fits do improve through an increase in scan rate as well as protein concentration. Also, as with the two-state irreversible model, the multimericity values calculated using the Lumry-Eyring model are 1.1 ± 0.2, i.e. very near unity, again suggesting no dissociation occurs as the protein is completely denatured. Interestingly, both of the non two-state models provided skewed ΔH, so that a large part of the total enthalpy is found in either the first or second step of the reaction, usually the latter. This finding lends additional credence to the two-state irreversible model as the best interpretation of our DSC data.

As the two-state irreversible model best describes the data, and the calculated multimericity is near unity, the two most likely models to describe the thermal denaturation of MR are found in schemes VII and VIII.

$$N_8 \xrightarrow{k} 8N' \xrightarrow{k} 8D \qquad \text{Scheme VII}$$

$$N_8 \xrightarrow{k} D_8 \qquad \text{Scheme VIII}$$

Scheme VII represents dissociation of the octameric enzyme into eight monomeric subunits, which subsequently proceed directly to the denatured state. This scenario surmises that the dissociation is not observed during, but before, the thermal denaturation of MR and that thermal denaturation of MR thus proceeds in the form of (eight) monomeric subunits. Thermal denaturation of MR was measured to occur

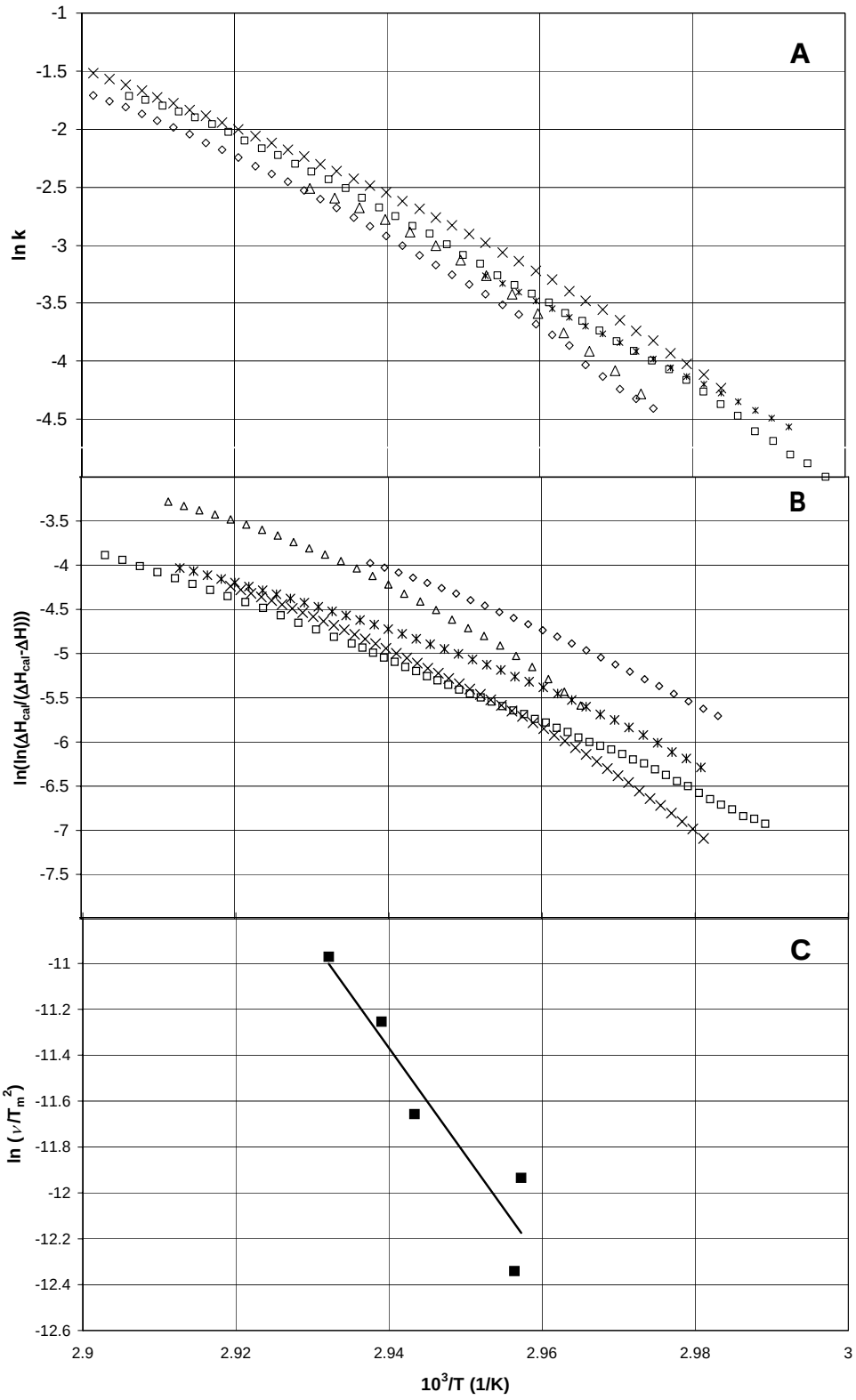


Figure 4-7. Methods used to determine the activation energy from experimental data. A: Arrhenius plot for kinetic constant, at scanning rates 0.5 K/min (□), 0.75 K/min (Δ), 1 K/min (∗), 1.5 K/min (×), 2 K/min(◊). B: The heat evolution versus 1/T, same scan rates. C: Scan rate dependence on maximum temperature.

between 60 and 68 °C. However, in previous work (*47*) we have determined the optimum temperature (T_{opt}) of reactivity for MR to be 50 °C, with very little activity at temperatures beyond 55 to 58 °C. Therefore, the enzyme possibly loses activity because of dissociation near 50 °C prior to unfolding at 60 and 68 °C. Catalytic activity of any other multimer of MR other than octamer has not been demonstrated. However, it has been shown that MR is a tetramer of dimers, with some of the active-site residues located at interacting points between the subunits (*1*). This finding is consistent with dissociation of the octamer near 50 °C and thermal transition around 60°C reflecting unfolding of the eight monomeric subunits.

Scheme VIII represents an alternative description of thermal denaturation of MR: the octameric enzyme is denatured prior to dissociation. Evidence for this scenario is the observation of aggregation of the protein at temperatures near the start of the transition, 58 °C.

4.5 Conclusion

Mandelate racemase is a tightly packed octameric enzyme, and it has been proposed that the structure is made up of a tetramer of dimers (*1*). Investigation of the thermal dissociation and unfolding of this enzyme could provide insight into the interactions of the subunits. DSC studies of the thermal denaturation of the MR revealed that the process is irreversible and kinetically controlled, and the transitions can be distorted

by aggregation. Through various studies, we have determined that the enzyme does not dissociate during thermal denaturation. After application of the experimental data to kinetic models it was determined that the consecutive two-step irreversible model and Lumry-Eyring pre-equilibrium models skew toward a two-state irreversible model, with the best fit observed using the two-state irreversible model.

4.6 REFERENCES

(1) Neidhart, D. J., Howell, P. L., Petsko, G. A., Powers, V. M., Li, R., Kenyon, G. L., and Gerlt, J. A. (1991) Mechanism of the Reaction Catalyzed by Mandelate Racemase. 2. Crystal Structure of Mandelate Racemase at 2.5 a Resolution: Identification of the Active Site and Possible Catalytic Residues. *Biochemistry 30*, 9264-9273.

(2) Milardi, D., La Rosa, C., and Grasso, D. (1996) Theoretical Basis for Differential Scanning Calorimetric Analysis of Multimeric Proteins. *Biophys. Chem. 62*, 95-108.

(3) Sanchez-Ruiz, J. M. (1992) Theoretical Analysis of Lumry-Eyring Models in Differential Scanning Calorimetry. *Biophys. J. 61*, 921-935.

(4) Michnik, A., Drzazga, Z., Kluczewska, A., and Michalik, K. (2005) Differential Scanning Microcalorimetry Study of the Thermal Denaturation of Haemoglobin. *Biophys. Chem. 118*, 93-101.

(5) Kenyon, G.l., Gerlt, J.A., Petsko, G.A., Kozarich, J.W. (1995) Mandelate Racemase: Structure-Function Studies of a Pseudosymmetric Enzyme. *Acc. Chem. Res. 28*, 178-186.

(6) Hasson, M.S., Schlichting, I., Moulai, J., Taylor, K., Barrett, W., Kenyon, G.l., Babbitt, P.C., Gerlt, J.A., Petsko, G.A., and Ringe, D. (1998) Evolution of an Enzyme Active Site: The Structure of a New Crystal Form of Muconate Lactonizing Enzyme Compared with Mandelate Racemase and Enolase. *Proc. Natl. Acad. Sci. 95*, 10396-10401.

(7) Kallarakal, A.T., Mitra, B., Kozarich, J.W., Gerlt, J.A., Clifton, J.G., Petsko, G.A., and Kenyon, G.l. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the K166r Mutant. *Biochemistry 34*, 2788-2797.

(8) Neidhart, D.J., Powers, V.M., Kenyon, G.l., Tsou, A.Y., Ransom, S.C., Gerlt, J.A, and Petsko, G.A. (1990) Preliminary X-Ray Data on Crystals of Mandelate Racemase. *J. of Biol. Chem. 263*, 9268-9270.

(9) Kenyon, G.l., and Hegeman, G.D. (1970) Mandelic Acid Racemase from *Pseudomonas Putida*. Evidence Favoring a Carbanion Intermediate in the Mechanism of Action. *Biochemistry 9*, 4036-4042.

(10) Bearne, S.l., and Wolfenden, R. (1997) Mandelate Racemase in Pieces: Effective Concentrations of Enzyme Functional Groups in the Transition State. *Biochemistry 36*, 1646-1656.

(11) Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent Evolution in the Enolase Superfamily: The Interplay of Mechanism and Specificity. *Arch. of Biochem. and Biophys. 433*, 59-70.

(12) Fee, J.A., Hegeman, G.D., and Kenyon, G.l. (1974) Mandelate Racemase from Pseudomonas Putida. Subunit Composition and Absolute Divalent Metal Ion Requirement. *Biochemistry 13*, 2528-2532.

(13) Gerlt, J. (2000) New Wine from Old Barrels. *Nat. Struct. Biol. 7*, 171-173.

(14) St. Maurice, M., and Bearne, S.L. (2002) Kinetics and Thermodynamics of Mandelate Racemase Catalysis. *Biochemistry 41*, 4048-4058.

(15) St. Maurice, M., Bearne, S.L. (2004) Hydrophobic Nature of the Active Site of Mandelate Racemase. *Biochemistry 43*, 2524-2532.

(16) Brewer, J. M., and Wampler, J. E. (2001) A Differential Scanning Calorimetric Study of the Effects of Metal Ions, Substrate/Product, Substrate Analogues and Chaotropic Anions on the Thermal Denaturation of Yeast Enolase I. *International J. of Biol. Macromolecules 38*, 213-218.

(17) Privalov, P. L. (1979) Stability of Proteins. Small Globular Proteins. *Adv. Prot. Chem. 33*, 167-241.

(18) Privalov, P. L. (1982) Stability of Proteins. Proteins Which Do Not Present a Single Cooperative System. *Adv. Prot. Chem. 35*, 1-104.

(19) Privalov, P. L. (1989) Thermodynamic Problems of Protein Structure. *Annu. Rev. Biophys. Chem. 18*, 47-69.

(20) Kurganov, B. I., Lyubarev, A. E., Sanchez-Ruiz, J. M., and Shnyrov, J. L. (1997) Analysis of Differential Scanning Calorimetry Data for Proteins: Criteria for Validity of One-Step Mechanism of Irreversible Protein Denaturation. *Biophys. Chem. 69*, 125-135.

(21) Conejero-Lara, F., Mateo, P. L., Aviles, F. X., and Sanchez-Ruiz, J. M. (1991) Effect of Zn^{2+} on the Thermal Denaturation of Carboxypeptidase B. *Biochemistry 30*, 2067-2072.

(22) Morin, P. E., Diggs, D., and Freire, E. (1990) Thermal Stability of Membrane-Reconstituted Yeast Cytochrome C Oxidase. *Biochemistry 29*, 781-788.

(23) Manly, S. P., Matthews, K. S., and Sturtevant, J. M. (1985) Thermal Denaturation of the Core Protein of *Lac* Repressor. *Biochemistry 24*, 3842-3846.

(24) Hernandez-Arana, A., Rojo-Dominguez, A., Altamirano, M.M. and Calcagno, M.L. (1993) Differential Scanning Calorimetry of the Irreversible Denaturation of *Escherchia Coli* Glucosamine-6-Phosphate Deaminase. *Biochemistry 32*, 3644-3648.

(25) Galisteo, M. L., Mateo, P. L., and Sanchez-Ruiz, J. M. (1991) Kinetic Study on the Irreversible Thermal Denaturation of Yeast Phosphoglycerate Kinase. *Biochemistry 30*, 2061-2066.

(26) Baker, D., and Agard, D. A. (1994) Kinetics Versus Thermodynamics in Protein Folding. *Biochemistry 33*, 7505-7509.

(27) Schimmele, B., and Pluckthun, A. (2005) Engineering Proteins for Stability and Efficient Folding, in *Protein Folding Handbook* (Buchner, J. a. T. K., Ed.) pp 1281-1333, Wiley-VCH Verlag GmbH & Co., Weinheim.

(28) Dill, K. A. (1990) Dominant Forces in Protein Folding. *Biochemistry 29*, 7133-7155.

(29) Sanchez-Ruiz, J. M. (1995) *Proteins: Structure, Function, and Engineering*, Vol. 24, Plenum, New York.

(30) Jelesarov, I., and Bosshard, H. R. (2004) Thermodynamics and kinetics of protein folding. Lecture notes: http://biocroma.unizh.ch/stabilitykinetics.pdf: pp 1-53.

(31) Hu, C. Q., and Sturtevant, J. M. (1989) A Differential Scanning Calorimetric Study on the Binding of Sulfate Ion and Cibacron Blue F3ga to Yeast Phosphoglycerate Kinase. *Biochemistry 38*, 813-818.

(32) Privalov, P. L., and Medved, L. V. (1982) Domains in the Fibrinogen Molecule. *J. Mol. Biol. 159*.

(33) Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Cortijo, M., and Mateo, P. L. (1988) Differential Scanning Calorimetry of the Irreversible Thermal Denaturation of Thermolysin. *Biochemistry 27*, 1648-1652.

(34) Sanchez-Ruiz, J. M., Lopez-Lacomba, J. L., Mateo, P. L., Vilanova, M., Serra, M. A., and Aviles, F. X. (1988) Analysis of the Thermal Unfolding of Procine Procarboxypeptidase a and Its Functional Pieces by Differential Scanning Calorimetry. *Eur. J. Biochem. 176*, 225-230.

(35) Freire, E., Van Osdol, W. W., Mayorga, O. L., and Sanchez-Ruiz, J. M. (1990) Calorimetrically Determined Dynamics of Complex Unfolding Transitions in Proteins. *Annu. Rev. Biophys. Biophys Chem. 19*, 159.

(36) Quesada-Soriano, I., Garcia-Maroto, F., and Garcia-Fuentes, L. (2006) Kinetic Study on the Irreversible Thermal Denaturation of

Schistosome Japonicum Glutathione S-Transferase. *Biochim. et Biophys. Acta 1764*, 979-984.

(37) Vogl, T., Jatzke, C., Hinz, H. J., Benz, J., and Huber, R. (1997) Thermodynamic Stability of Annexin V E17g: Equilibrium Parameters from an Irreversible Unfolding Reaction. *Biochemistry 36*, 1657-1668.

(38) Thorolfsson, M., Ibarra-Molero, B., Fojan, P., Peterson, S. B., Sanchez-Ruiz, J. M., and Martine, A. (2002) L-Phenylalanine Binding and Domain Organization in Human Phenylalanine Hydroxylase: A Differential Scanning Calorimetry Study. *Biochemistry 41*, 7573-7585.

(39) Lyubarev, A. E., Kurganov, B. I., Burlakova, A. A., and Orlov, V. N. (1998) Irreversible Thermal Denaturation of Uridine Phosphorylase from *Escherichia Coli* K-12. *Biophys. Chem. 70*, 247-257.

(40) Guzman-Casado, M., Parody-Morreale, P. I., Mateo, J. M., and Sanchez-Ruiz, J. M. (1990) *Eur. J. Biochem. 188*, 181.

(41) Lepock, J. R., Rodahl, A. M., Zhang, C., Heynen, M. L., Waters, B., and Cheng, K. H. (1990) *Biochemistry 29*, 681.

(42) Lyubarev, A. E., and Kurganov, B. I. (2000) Analysis of Dsc Data Relating to Proteins Undergoing Irreversible Thermal Denaturation. *J. of Thermal Analysis and Calorimetry 62*, 51-62.

(43) Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate Racemase and Muconate Lactonizing Enzyme Are Mechanistically Distinct and Structurally Homologous. *Nature 347*, 692-694.

(44) Gander, W., and Gautschi, W. (2000) Adaptive Quadrature-Revisited. *BIT-Numerical Methods 40*, 84-101.

(45) Dormand, J. R., and Prince, P. J. (1980) A Family of Embedded Runge Kutta Formulae. *J. Comp. Appl. Math. 6*, 19-26.

(46) Takahashi, K., and Sturtevant, M. M. (1981) Thermal Denaturation of Streptomyces Subtilisin Bpn' and the Inhibitor-Subtilisin Complex. *Biochemistry 20*, 6185.

(47) Thaler, T. L., Gibbs, P. R., Trebino, R. P., and Bommarius, A. S. (2006) Search for Extraterrestrial Life Using Chiral Molecules: Mandelate Racemase as a Test Case. *Astrobiology 6*, 901-910.

CHAPTER 5

INVESTIGATION OF MR REACTIVITY AT LOW TEMPERATURES AND IN NON-AQUEOUS MEDIA

The planets and moons focused on in the search for extraterrestrial life are at temperatures much lower than those found on Earth. Since each step of the catalytic cycle is sensitive to changes in temperature it is of interest to investigate the enzymatic reaction towards temperatures observed on Mars (-40°C). Investigation of enzyme reactivity at low temperatures requires a variety of changes to methods commonly used at ambient conditions. Enzymatic rate measurements at low temperatures cannot be conducted in a purely aqueous environment; this renders it necessary to perform the reaction in a medium with a melting point below the temperatures in question. Investigation of enzyme activity in non-aqueous media is also important as most extraterrestrial environments contain only small amounts of water. As life on Earth has evolved to require water as a solvent, enzymes lose activity in organic solvents. The goal of this Aim is to explore the activity of MR in low temperature conditions as well as investigate MR activity in low-water environments.

Organic cryosolvents are the most common medium used for measurement of enzyme reactions at temperatures less than 0°C; however, such solvents tend either to denature the protein or vastly increase the K_M value and, therefore, are not favorable in the enzymatic environment (*1*). Another option is the use of ***highly concentrated aqueous salt solutions.*** Saturated ammonium salt has been used successfully as a cryosolvent (*2*) to measure β-lactamase activity in temperatures as low as -60°C. However, the water content needs to be less than 20% to prevent phase separation below -120°C (*3*).

A third cryosolvent system that provides an environment to measure enzymatic activity at low temperatures is the use of ***water-in-oil microemulsions*** which encapsulate the enzyme and substrate. Water-in-oil microemulsions, tiny water droplets stabilized by a surfactant in an organic solvent continuum, have been used to mimic the cell membrane (*4-7*) as well as act as micromembranes for enzyme catalysts (*8-11*). These microemulsions create a sheltered environment for the enzyme.

Cryosolvents composed of small molecules such as ammonia, formic acid, and lower hydrocarbons (such as methane or ethane), are ubiquitous in the Universe and have been found on many bodies in the Solar System. For example, the Huygens probe has identified dark patches that resemble pools of liquid on Titan that scientist expect to be responsible for the high concentrations of hydrocarbons, mostly methane

and ethane (*12*, *13*). Thus, it is possible for concentrated ammonium salt solutions and water-in-oil microemulsions to form on moons and planets as long as a small amount of liquid water is present. In the case of microemulsions, the additional presence of any molecule that can act as a surfactant is also required. In this work, we have employed concentrated ammonium salts and water-in-oil microemulsions as well as organic cryosolvents to investigate MR activity at low temperatures as well as in low aqueous environments.

5.1 Background

5.1.1 Enzymatic reactions at low temperatures

Studies of enzymatic reactions at low temperatures can provide valuable information about the enzyme. Douzou (*14*) began using cryoenzymology in the late 1970's for stabilization of enzyme-substrate intermediates at low temperatures. By slowing down the kinetics of their formation, intermediates could be studied without changing the mechanism. This trapping of intermediates in turn resulted in a method for elucidating the enzyme mechanism (*14*).

Low-temperatures studies also allow for study of function because of decreased rate. The molecular interactions are slowed, rendering it easier to determine the structure and dynamics of the enzyme (*3*).

Another application of enzymatic low temperature studies is the storage of food or pharmaceuticals in frozen or liquid solutions (*15*).

5.1.1.1 Temperature Effects

It is important to determine the effects low temperatures will have on the enzymatic reaction. First, it has been suggested that most proteins undergo a glass transition (temperature at which molecules have very little mobility) around -50°C. At this point the structure becomes rigid, and thus substrate binding may be prevented (*16*), in turn, the enzyme's function at this transition temperature can be extensively altered. However, the ability of enzymes to catalyze reactions does not stop altogether, at least in some cases. For example, enzymatic activity has been measured at lower temperatures, with Bragger et. al. having measured activity of β-lactamase down to -100°C with no apparent lower limit (*17*).

Low temperatures could also cause a conformational change in the enzyme. Temperature-induced dissociation has been shown in oligomers, such as lactate dehydrogenase (*18*). Molecular flexibility is a tradeoff between conformational stabilility and enzyme activity. The activity of an enzyme depends on its flexibility, but a less flexible enzyme is more stable. For example, it has been found that enzymes from thermophiles are more stable i.e. less flexible, but also have lower activity than those from mesophiles, however, this tradeoff allows for the

organisms survival at elevated temperatures (*19*). Even more interesting, psychrophilic enzymes have been found to be even more flexible (*19*). These cold adapted enzymes have a Gibbs Free enthalpy of activation ($\Delta G^{\neq}$) much below that found in mesophiles. This low activation free enthalpy is due to a decrease in the activation enthalpy, indicating that conformational changes needed to reach the activation state are decreased (*19*). The extent of temperature effects is governed by the Arrhenius activation energy (E_a). It has been found in some cases that the Arrhenius activation energy can be raised by cryosolvents (*20*), which in turn raises the rate at which the activity decreases with decreasing temperature (*17*). There is however a caveat in that the use of some cryosolvents can have structural effects on the enzyme that lead to unfolding of the protein, decreasing the enzymatic activity (*16*, *20*).

Temperature dependence is commonly described by the Arrhenius equation (Equation 5-1).

$$k = Ae^{E_a/RT} \qquad 5\text{-}1$$

where: k is the rate constant, A is the pre-exponential factor, E_a is the activation energy, R is the gas constant, and T is the absolute temperature (in K). A general rule of thumb for most chemical and biological systems is that the rate of a chemical reaction is lowered by a factor of 2 to 3 with every 10°C decrease in temperature.

5.1.2 Biocatalyst in non-aqueous media

Working at low temperatures requires some sort of antifreeze, or non-aqueous media. There are advantages in using biocatalysts in non-aqueous media, including increased solubility of non-polar substrates and products, suppression of water-dependent side-reactions, and prevention of microbial contamination (*21*). Increasing the activity of an enzyme in non-aqueous environments has a variety of other applications including biocatalytic membranes, coatings, and paints, biocatalytic plastics, and biocatalytic polymer composites, and synthesis of fine chemicals and pharmaceuticals (*22*).

However, enzymatic activity is typically decreased or completely eliminated in low-water environments (*23*). Some contributing factors include the loss of necessary water molecules at the surface of the enzyme as well as a decrease in polarity of the enzyme's environment (*24*). Many strategies, such as protein engineering and water-in-oil microemulsions have been employed to improve the enzyme stability and activity (*25*).

5.1.2.1 Organic solvents

Choosing an organic solvent as a low-water environment for biocatalysis is not always the best choice, because in many cases for the organic solvent to be effective high concentrations at greater than 40% are necessary (*16*). Also, many times the enzyme is not soluble in the

organic solvent, except for the use of protic organic solvents such as DMSO, ethylene glycol and formamide, which can also inactivate the protein (*25*). Not only will these high concentrations affect the bulk properties of the solvent, they also affect the protein structure. This is because the organic solvent increases the activity coefficient of the charged amino acids on the surface of the protein, which can favor precipitation. In contrast, the solvent will decrease the activity coefficient of the uncharged amino acids, exposing them to the solvent, in turn unfolding the protein (*26*). However, some solvents can be stabilizing such as those with multiple hydroxyl groups (*16*). It has also been determined that the K_M value can greatly increase in organic solvents (*2*).

There have been attempts to correlate activity with physicochemical properties. A general rule was proposed using the partition coefficient (log P) in relation to the enzymes compatibility with the organic solvent. If the organic solvent has a log P greater than 4 it is most likely innocuous to the biocatalysts, however, if the log P is less than 2 the organic solvent will lead to deactivation of the enzyme. Conditions at medium polarity in between are more difficult to estimate (*27*).

The main reason enzymes lose activity in organic solvents is the unfolding of the enzyme. It has been suggested that the cause is exposure in organic solvents of the protein hydrophobic regions, normally buried.

5.1.2.2 High concentration salt solutions

Cartwright and Waley (1987) chose a high concentration of salt solutions (specifically ammonium acetate), which allowed them to measure β-lactamase activity down to –60 °C. They found that the K_M value was not raised, making this a more desirable cryosolvent than some of the organic solvent systems (*2*).

5.1.2.2.1 Protein deactivation

Using a high concentration of salt with an enzyme could be dangerous because dissolved salts affect properties of proteins in solution including melting temperature. Therefore, it is necessary to further investigate how these systems affect the stability of the enzyme. At moderate to high concentrations of salt, the electrostatic shielding effects are not effective (*28*), increasing the possibility for enzyme deactivation.

The lower water activity indicates a less ideal solution caused by the increased salt content. Broering and Bommarius (2005) (*28*) have described a model for prediction of salt effects in the Hofmeister series on deactivation kinetics. They found that it is possible to predict rate constants of deactivation in chaotropic solutions using deactivation constants and B-viscosity coefficients. The B-viscosity coefficient is a value that characterizes ion-solvent interactions at a given temperature and pressure (*29*). They also determined that there is a lower limit to deactivation rate constants in kosmotropic solutions, suggesting that such

salt solutions are preferred to those containing chaotropes for enzyme reaction measurements.

5.1.2.3 Water-in-oil microemulsions

Another cryosolvent system that can be used to create a low-water environment is water-in-oil microemulsions, which encapsulate the enzyme and substrate protecting them from the organic solvent environment. Water-in-oil microemulsions, tiny water droplets stabilized by a surfactant in an organic solvent continuum, have been used to mimic cell membrane (*5*) as well as act as micromembranes for enzyme catalysts (*8-11*). The microemulsions consist of non-polar tails of surfactant molecules in contact with an organic phase, while the polar heads sequester together forming an enclosed water vesicle. These systems have been used to mimic many life systems (*5*).

The size and shape of a water-in-oil microemulsion is defined by w_o which is the molar ratio of the water to surfactant. These optically transparent, thermodynamically stable systems are formed under high surfactant concentrations in non-polar solvents (*16*). Enzymatic activity has been detected in many enzymes such as lipases, serine proteases, lysozyme, alcohol dehydrogenase as well as others (*30*).

5.1.3 MR and organic solvents

Reactivity of mandelate racemase in the presence of organic solvents has not been promising. In 2002, the Faber group tested the log P rule for MR and found that for log P values greater than zero there was a slight enhancement of activity. They found MR is somewhat inert toward polar solvents and totally inactive in organic solvents with a log P greater than or equal to one (*21*). MR has also been tested in both organic solvents and ionic liquids as a new medium for a dynamic resolution approach. In ionic liquids, it was found that there was a loss of enzymatic activity in a water activity less than 0.8 (*31*).

The compatibility of MR with organic solvents has been explored by Pogorevc et. al. (*21*), who exposed the enzyme to a range of neat organic solvents with a wide range of polarity and both water-miscible and non-miscible solvents. Incorporating their results, we tried other systems more representative of a possible composition on Mars, Europa, or Titan, such as concentrated ammonium salt solutions and water-in-oil microemulsions.

5.2 Materials and Methods

5.2.1 Materials

(*R*)-mandelic acid was a kind gift of BASF (Ludwigshafen Germany); (*S*)-mandelic acid, citric acid, formic acid, ammonium iodide, ammonium bromide, ammonium chloride, ammonium sulfate, ammonium nitrate and

Tris(hydroxymethyl)-aminomethane, sodium bicarbonate, bistris propane, and Triton-X100 UltraSigma were obtained from Sigma-Aldrich (St. Louis, Missouri), $MgCl_2$ was from Fisher Chemical (Fair Lawn, New Jersey), Luria Broth, Aerosol OT, ammonium acetate, heptane, methanol, and sodium phosphate from EM Science (Gibbstown, New Jersey), ampicillin from Shelton Scientific (Shelton, Connecticut), ammonium formate and 1-hexanol from Alfa Aesar (Pelham, New Hampshire), ethylene glycol, ethanediol, acetone, methoxyethanol, dimethylformamide (DMF), dimethylsulfoxide (DMSO) were obtained from JT Baker (Phillipsburg, New Jersey).

5.2.2 Organic solvent systems

Organic solvent percentages are by volume. The substrate solution was made to a final concentration of 20 mM. Optical rotation of the organic solvent mixture with buffer, minus the substrate was measured to subtract out any background. Enzyme at a final concentration of 0.5 mg was added to the organic solvent/substrate solution and optical rotation was measured at room temperature for up to two hours. Those showing no activity were incubated at 23 °C for 24 hours to verify no change in optical rotation.

5.2.3 Concentrated salt solutions

Ammonium acetate and ammonium formate cryosolvents were prepared according to Cartwright and Waley (*2*). All measurements were made at 50% saturated solution and pH 7.6 (in reference to 4 °C, at which the salt solution was prepared). The enzyme solution was prepared in 20% saturated salt solution. Enzyme activity was measured at different temperatures employing an endpoint method. Each reaction was carried out in 8 mL substrate in 50% saturated salt solution (with a final concentration of 100 mM (*R*)-mandelic acid, 50 mM Tris, and 1 mM $MgCl_2$). The reactions were stopped at the desired time point by the addition of 50 mM citric acid (pH 3.0). The optical rotation of the solution was then measured after the solution reached a temperature of 24°C. The solutions were allowed to mix for 4 minutes, the time for introduced bubbles to disperse. For temperatures greater than 0°C measurements were taken every minute for six minutes, for temperatures less than 0°C measurements were taken every five minutes for thirty minutes.

5.2.3.1 Salt Deactivation experiments

Salt deactivation experiments were carried out in salt solutions with a constant water activity (a_w) of 0.97. Salt concentrations for this water activity were calculated using the computer program "Aqueous Solutions," designed by the IUPAC (*32*). The enzyme was dissolved in 50

mM Hepes (pH 7.5) and then mixed with an equal amount of salt buffered in 50 mM Hepes (pH 7.5) to give the desired salt concentration. Protein concentrations were measured using Coomassie stain (Pierce, Rockford, IL) in a Bradford assay, with a final enzyme concentration being 1 mg/mL.

Enzyme salt solutions were placed in 0.65 mL thin-walled tubes and placed into a water bath at the deactivation temperature, which was 55°C. One tube was kept on ice for a time zero measurement. Tubes then were removed at different times depending on the salt used. The tubes were placed in ice water immediately after removal from the water bath to stop the deactivation reaction.

The tubes were centrifuged at 13000 xg for 5 minutes to collect the precipitate, which could interfere with enzymatic assay. The supernatant (10 μL) was assayed, according to the polarimetry assay as previously described, for residual activity (relative to the undeactivated sample) The first order deactivation constants ($k_{1,obs}$) were calculated from the slopes of ln (residual activity) versus deactivation time plots.

5.2.4 Water-in-oil microemulsion systems

Two types of microemulsion systems were prepared: i) an anionic surfactant system of 0.2 M AOT surfactant in heptane with 10 mM mandelate in 0.1 M phosphate buffer at pH 7.5, and ii) a non-ionic surfactant system of 0.2 M Triton-X 100 (Poly(oxyethylene) p-tert-octylphenyl ether) in a 4:1 n-heptane/n-hexanol mixture as organic

continuum and 16 mM mandelate in 0.2 M sodium bicarbonate buffer at pH 7.0. The water-surfactant ratio w_0, which scales linearly with the radius of the microemulsions (*16*), was set at w_0 = 10 for the AOT system, the lowest possible given the sensitivity of our polarimeter, corresponding to a freezing point of -18 °C (*1*), and at w_0 = 15 for the Triton system, at its maximum of enzymatic activity.

Each temperature reaction was carried out in 8 mL. The microemulsions were first prepared using the w_o values determined above (with a final concentration of 285 mM (*R*)-mandelic acid, 50 mM Tris, and 1 mM $MgCl_2$). The solution was cooled or heated to the desired temperature and then the enzyme was added to the solution and vortexed to make sure it was well mixed. The reactions were stopped at the desired time point by the addition of 50 mM citric acid (pH 3.0). The optical rotation of the solution was then measured after the solution reached a temperature of 24°C. For temperatures greater than 0°C measurements were taken every minute for six minutes, for temperatures less than 0°C measurements were taken every five minutes for thirty minutes.

5.3 Results

5.3.1 MR in organic solvents

There are several water-solvent systems that have been used to measure MR reactivity (*33*, *34*) . In general, the most common organic

solvent system for low temperature studies is 70% methanol/10% ethylene glycol (EG) (*23*). However, we did not observe activity for the MR reaction past 50% MeOH/ 10% EG (Figure 5-1). Increasing the amount of ethylene glycol to 20% (results not shown) in hopes of stabilizing the enzyme again resulted in no MR activity, this time past 30% MeOH.

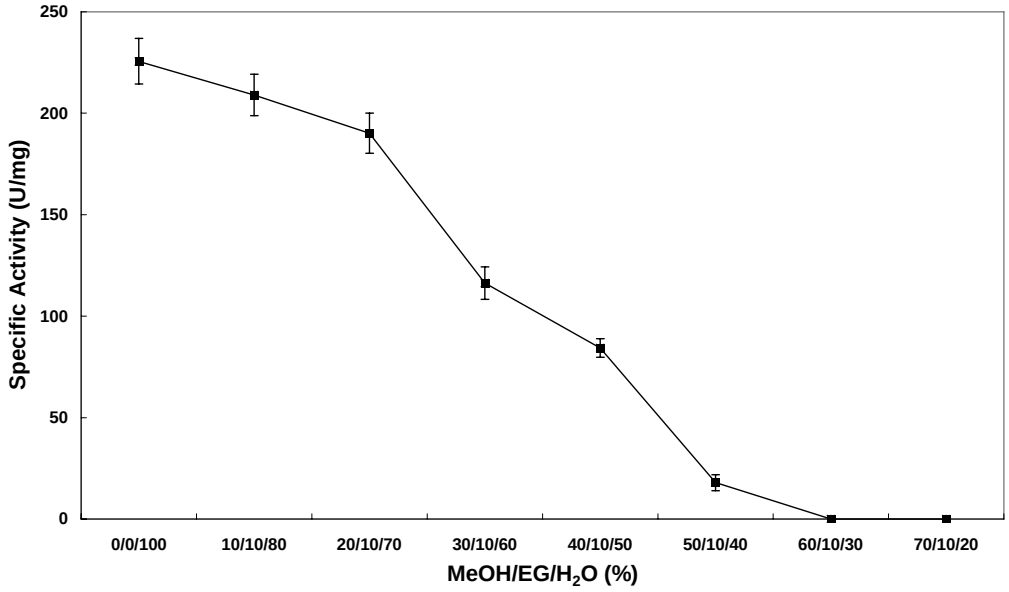


Figure 5-1. MR activity in varying concentrations of methanol/ethylene glycol/H_20.

Next, we tried DMSO/EG/Substrate, we did not observe MR activity beyond 10% DMSO (Figure 5-2). Although we did observe some activity at low concentrations of organic solvent, for the measurement of the reaction at low temperature we need the organic solvent systems at higher concentrations. In order to be used in low temperature

investigations, the water concentration cannot exceed 20% (*1*); therefore there has to be a high concentration of organic solvent. We investigated more solvent systems and found no activity in any except methoxyethanol:ethanediol, which was very small (about 10 U/mg) (Table 5-1).

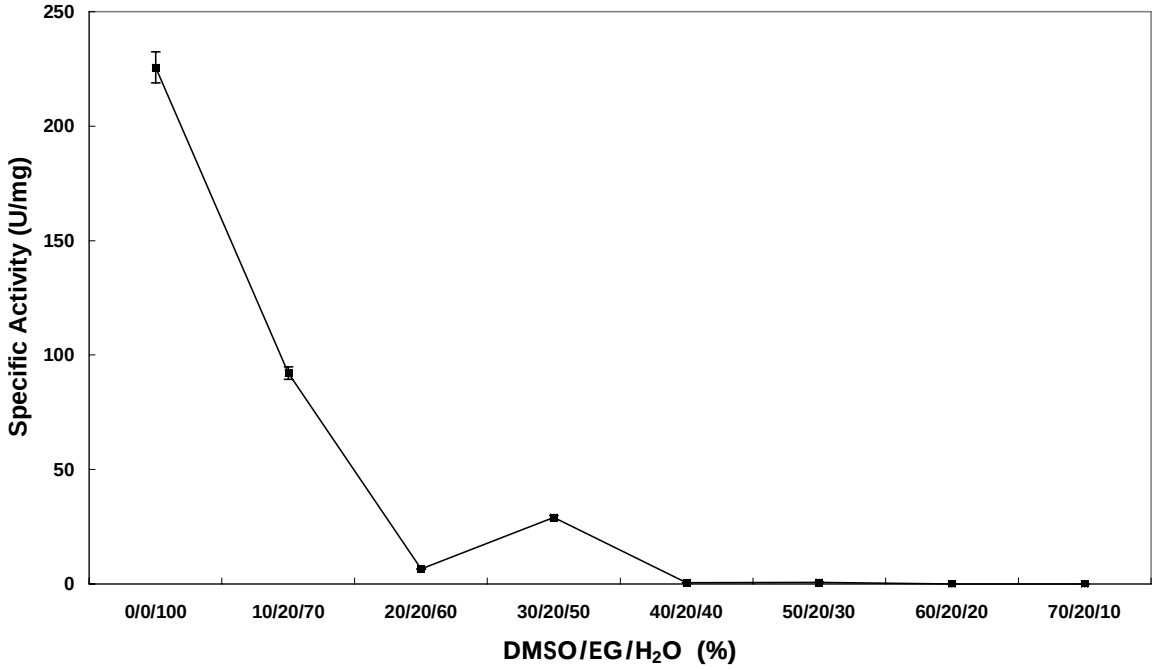


Figure 5-2. MR activity in varying concentrations of DMSO/EG/H_2O.

Table 5-1. Organic solvent systems used as a non-aqueous media for MR reaction.

Solvent System	MR Activity (U/mg)-90 mM mandelic acid
Methanol:ethandiol (70:10)	0
Methanol:glycerol (70:10)	0
Acetone:ethanediol (50:30)	0
Acetone:methoxyethanol:ethanediol (35:35:10)	0
DMF:ethanediol (70:10)	0
DMF (80)	0
Methoxyethanol (80)	0
Methoxyethanol:ethanediol (70:10)	Slight; about 10 U/mg

5.3.2 MR in concentrated ammonium salts

First, we measured specific activity of MR in both ammonium salt solutions at 24°C as a function of salt saturation (degree of saturation determined at 4°C), both salts behaved similarly (Figure 5-3). Activity

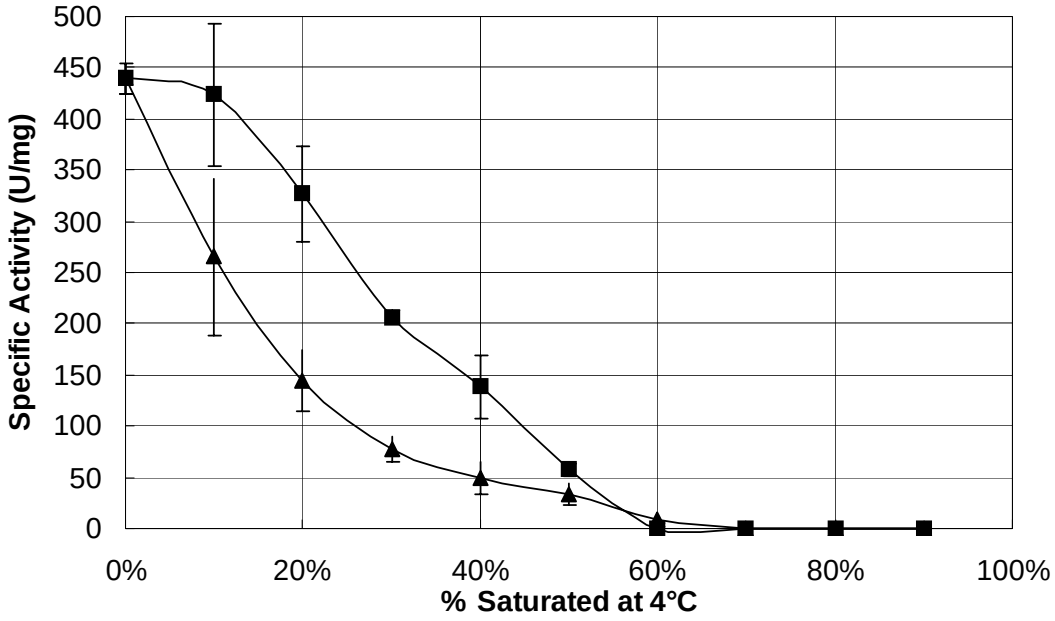


Figure 5-3. Measurement of MR specific activity at varying concentrations of ammonium acetate (+) and ammonium formate (▲).

decreased linearly for ammonium acetate, after a lag distance of about 10% and negatively exponentially for ammonium formate. We determined 50% saturation as the maximum salt content with measurable MR activity.

Following reports of investigations of enzyme activity in sub-zero temperatures by Cartwright and Waley (1987) (*2*), we prepared ammonium salt solutions saturated (pH 7.6) to different degrees at 4°C with 100 mM mandelate (pH 7.5) and determined their freezing points. The freezing points of both ammonium acetate/ and ammonium formate/mandelate solutions superimpose at all salt concentrations and decrease linearly with degree of saturation (Figure 5-4).

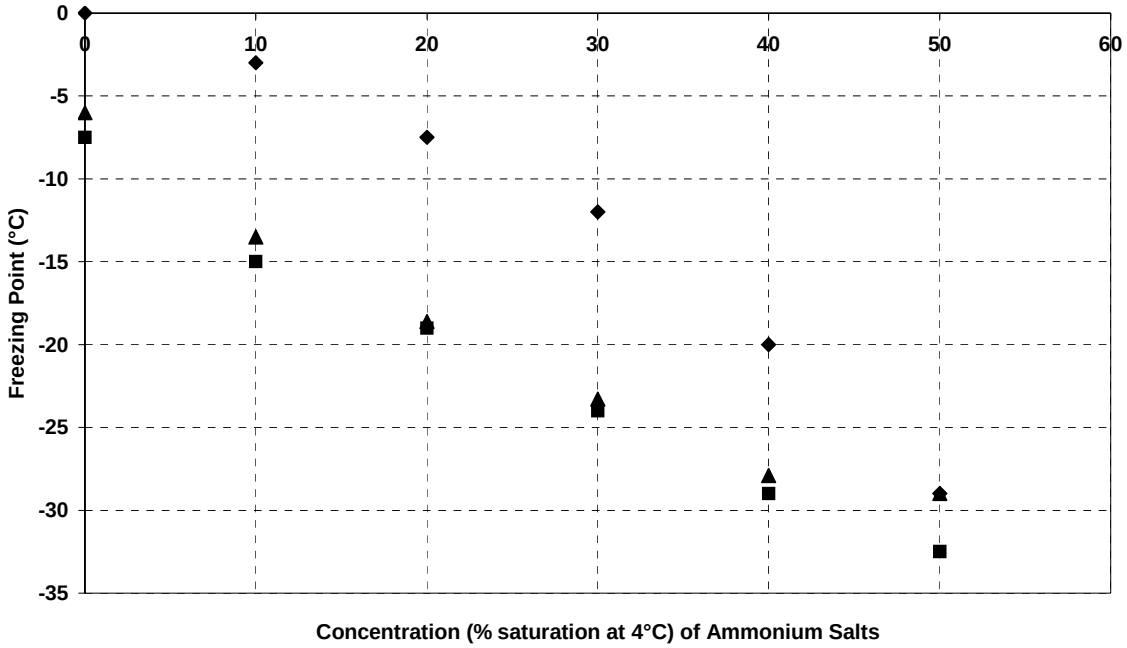


Figure 5-4. Freezing point of concentrated ammonium salts. Ammonium acetate (+), ammonium formate (▲), and ammonium acetate measured by Cartwright and Waley (1987) (♦).

We found that the solutions began forming ice crystals around -30°C, with ammonium acetate at –32°C and ammonium formate at –29°C. Ammonium bicarbonate did not yield a significant freezing point depression. Next, we determined the specific activity of MR in 50% saturated ammonium salt solution as a function of temperature over the largest possible temperature range, from -30 to 70 °C. Below –30 °C, specific activity was too small and both salt systems tended to freeze. The optimum temperature for MR activity in both salt systems is at 60 °C, whereas at 70°C, activity quickly decreases (Figure 5-5). Both Arrhenius (Figure 5-6) and Eyring plots (Figure 5-7) were created from this data and activation parameters were calculated (Table 5-2).

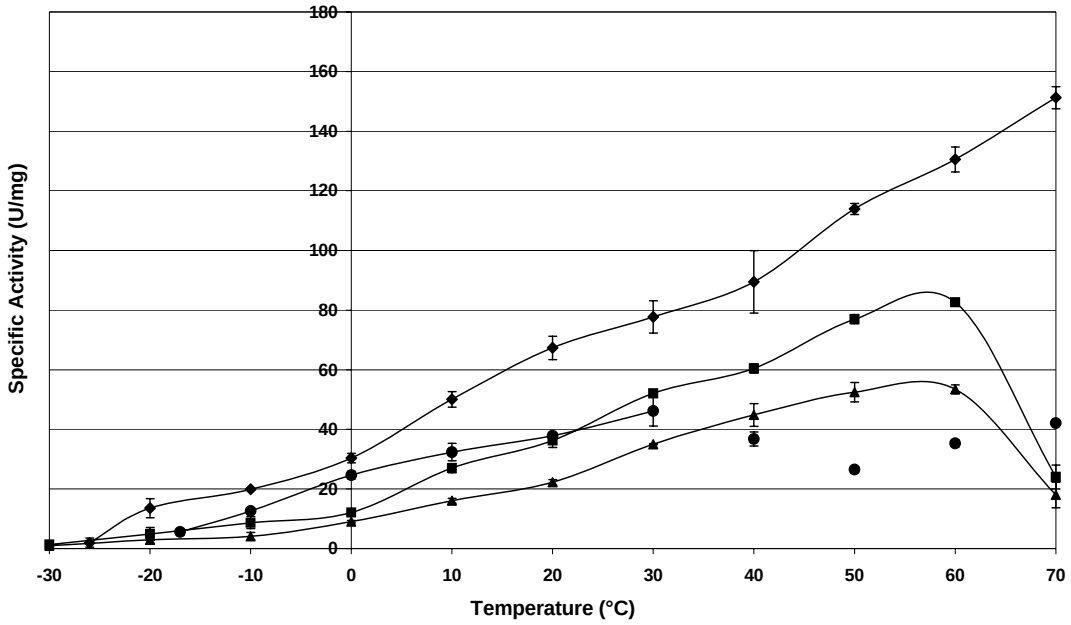


Figure 5-5. Specific activity over temperature for ammonium acetate (■), ammonium formate (▲), AOT-phosphate system (●), and Triton-bistris propane system (♦).

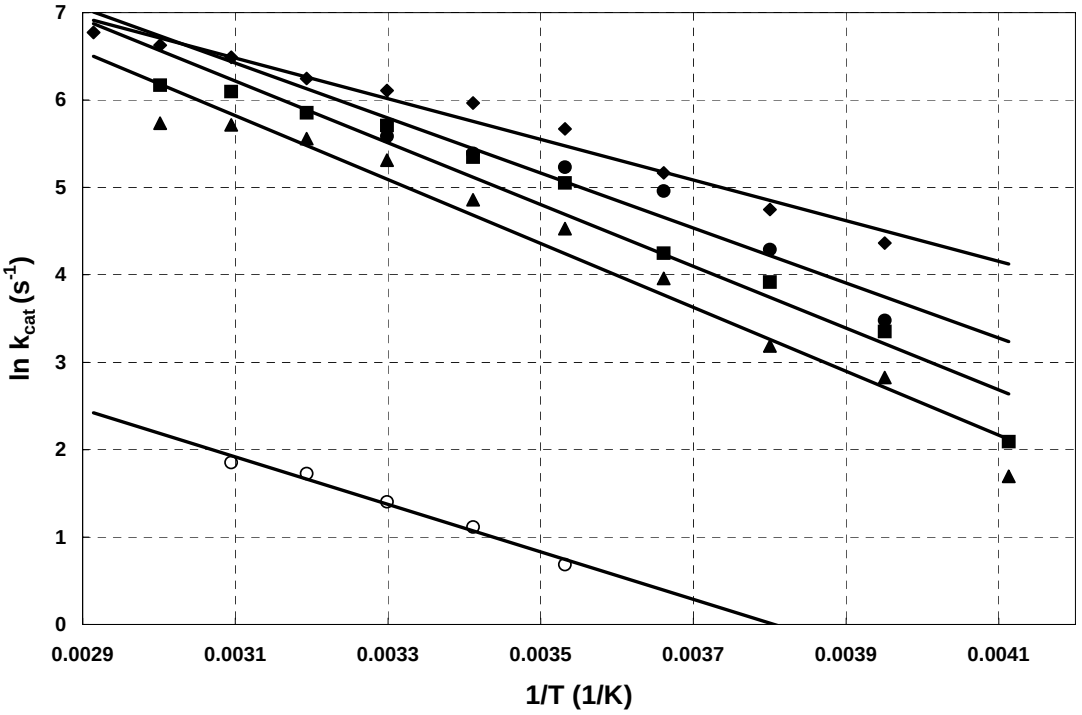


Figure 5-6. Arrhenius plot of MR in concentrated ammonium salts, and water-in-oil microemulsions. Tris (O), ammonium acetate (■), ammonium formate (?), AOT-phosphate system (•), and Triton-bistris propane system (♦).

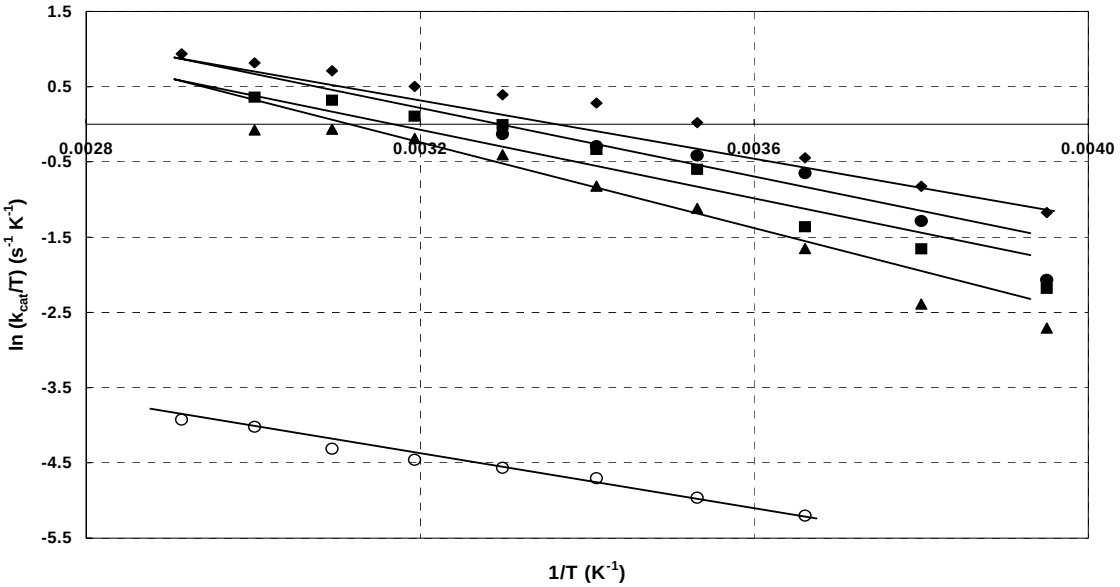


Figure 5-7. Eyring plot of MR in concentrated ammonium salts, and water-in-oil microemulsions. Tris (O), ammonium acetate (■), ammonium formate (?), AOT-phosphate system (•), and Triton-bistris propane system (♦).

Table 5-2. Activation parameters of mandelate racemase (MR) from *Ps. putida* in aqueous buffer, concentrated ammonium salt solutions, and reversed micellar systems.

	E_a (kJ/mol)	$\Delta H^{\neq}$ (kJ/mol)	$\Delta S^{\neq}$ (J/mol K)	$\Delta G^{\neq}$ (kJ/mol)
MR aqu. buffer[a]	26.4 ± 1.3	23.9 ± 1.2	-118.2 ± 20	54.2 ± 2
MR HEPES (lit)[c]	NR	64.5 ± 2	8.37 ± 0.4	56.1 ± 2
MR PO_4 (lit)[d]	NR	133.6 ± 4	-38.9 ± 2	146.5 ± 9
MR NH_4CH_3COO	29.4 ± 1.8	27.1 ± 1.4	-197.0 ± 19	81.0 ± 4
MR NH_4HCOO	30.4 ± 1.2	28.0 ± 1.4	-197.0 ± 28	82.0 ± 4
MR Triton	19.3 ± 0.6	16.9 ± 0.7	-139.0 ± 22	55.0 ± 2.2
MR AOT	27.2 ± 3	24.8 ± 3.2	-115.0 ± 15	56.2 ± 7.3

Legend: aqu. buffer: 0.05 M Tris, pH 7.5, NH_4CH_3COO: ammonium acetate, NH_4HCOO: ammonium formate, Reversed micellar systems: Triton: 0.2 M Triton X-100/20% n-hexanol/80% n-heptane/$NaHCO_3$, w_0 = 15; AOT: 0.2 M Aerosol OT/n-heptane/PO_4 at w_0 = 10; Lit. reference: c: 0.2 M HEPES, pH 7.5 d: 0.1 M phosphate, pH 7.5
NR: Not Reported.

5.2.4.1 Deactivation of MR in concentrated ammonium salts

Using the "Aqueous Solutions" programs we determined the salt concentrations for a water activity of 0.97 (Figure 5-8). There was no data for ammonium formate or ammonium acetate, knowing where these should be on the plot we estimated these values to be 0.97 mol/kg and 1.05 mol/kg. We measured observable deactivation kinetics of MR at 50°C and a_w of 0.97 in 50 mM HEPES (pH 7.5) for the following salts: NH_4I, NH_4Br, NH_4Cl, NH_4HCOO, NH_4CH_3COO, $(NH_4)_2SO_4$, and $(NH_4)_2NO_3$. A plot of residual activity versus time was made for each salt, and the deactivation constants were calculated from the slope of the plot, an

example of this is shown in Figure 5-9. To predict the rate constant of deactivation, we plotted ln of the deactivation constant versus the Jones-Dole B-viscosity coefficient (Figure 5-10).

We determined the order of deactivation and compare this result to other proteins such as horse liver alcohol dehydrogenase (HL-ADH), a-chymotrypsin, and monomeric Red Fluorescent protein (Table 5-3). These enzymes have been previously used to develop a model for prediction of salt effects on enzymes (*28*).

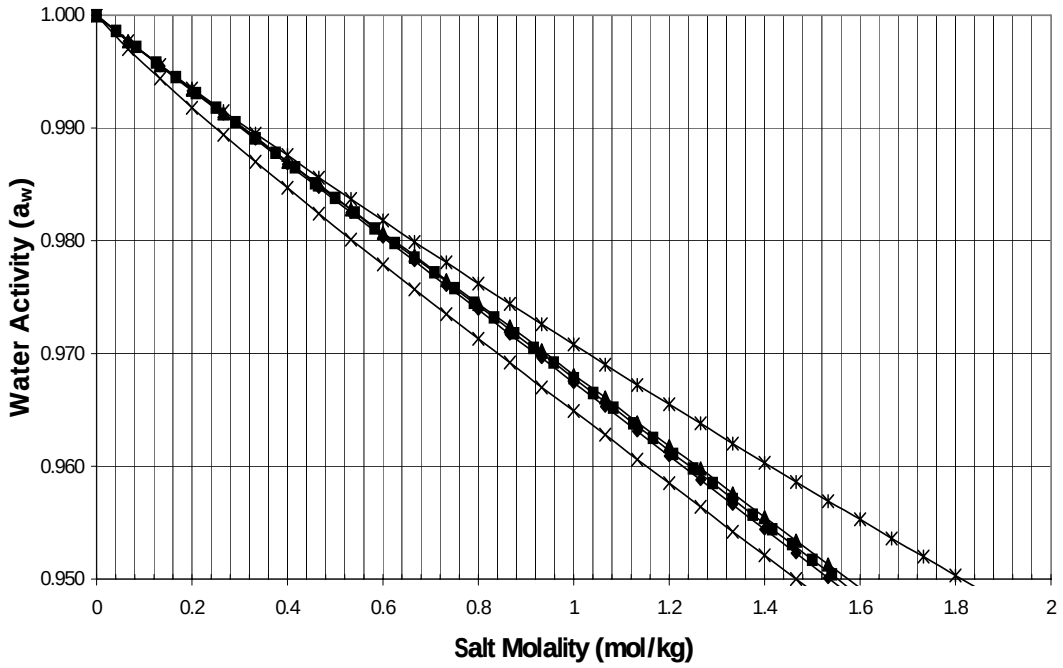


Figure 5-8. Measurement of salt concentrations at aw = 0.97, using osmotic coefficients. Ammonium iodide (♦), ammonium bromide (■), ammonium chloride(▲), ammonium sulfate (×), and ammonium nitrate (∗).

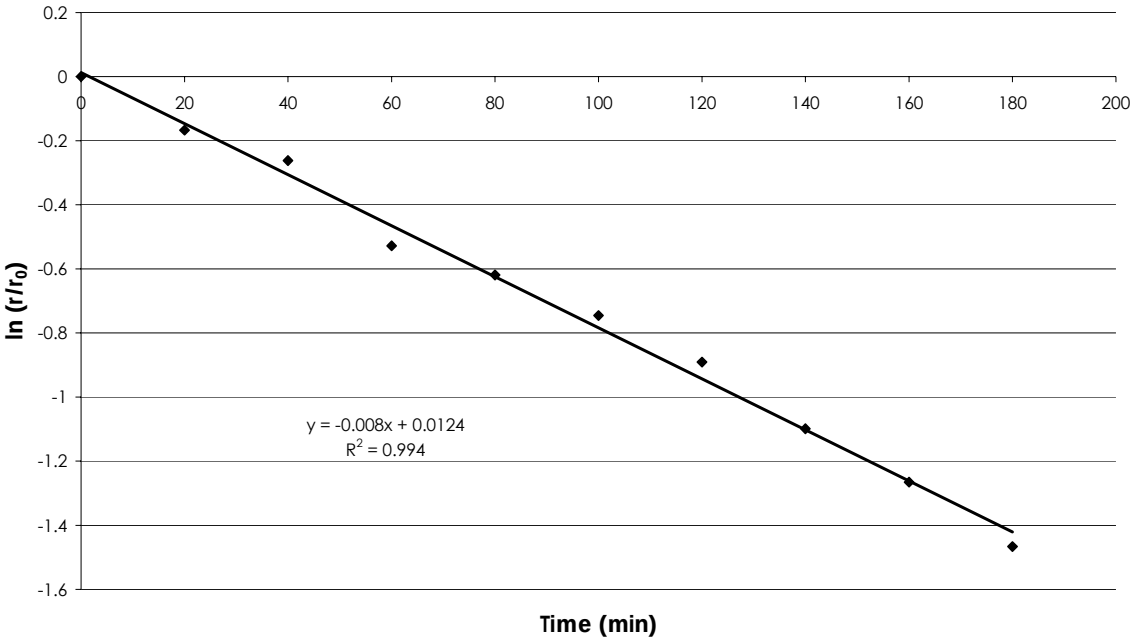


Figure 5-9. Example of raw deactivation data for mandelate racemase (1 mg/ml) in $(NH_4)_2NO_3$.

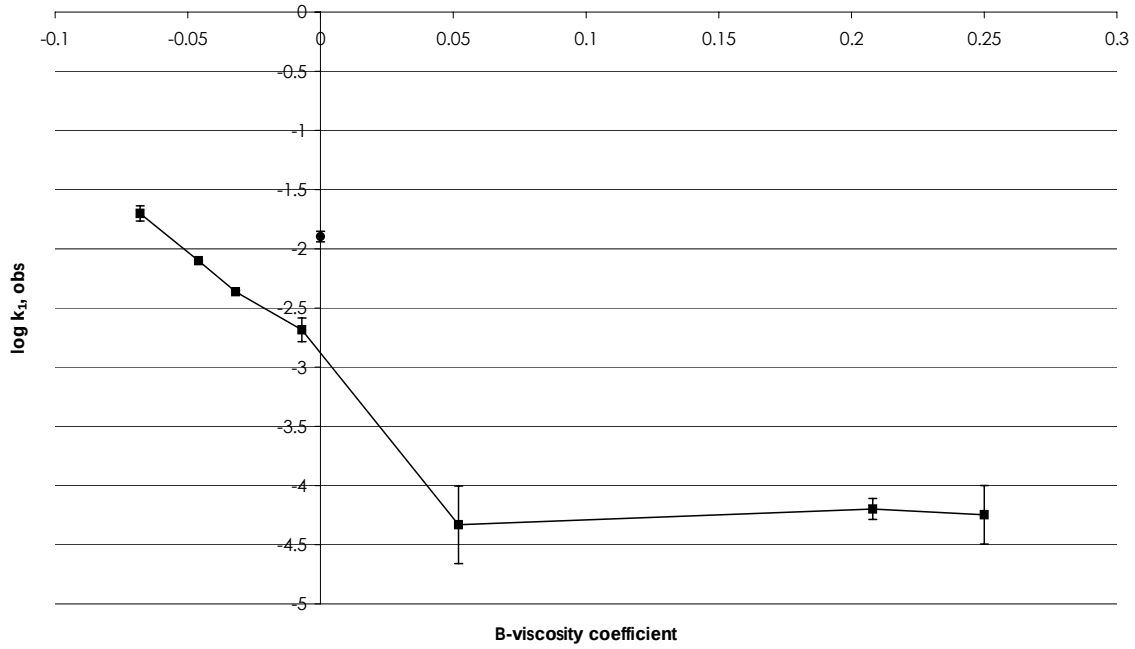


Figure 5-10. The logarithm of the observed deactivation constant, k1, obs versus the Jones-Dole B-viscosity coefficient of the anion.

Table 5-3. Comparison of salt deactivation data.

Enzyme	Chaotrope slope	Kosmotropic slope
Monomeric Red Fluorescent protein	-13.0	-2.647
Horse liver alcohol dehydrogenase	-35.5	-3.075
Formate dehydrogenase	Not Reported	-4.266
Mandelate racemase	-16.2	-3.075

5.3.3 MR in water-in-oil microemulsions

Lack of enzymatic activity in the organic solvent systems led us to explore the use of water-in-oil microemulsions. After testing a variet of buffers with varying enzyme concentrations, we chose two different systems: AOT/n-heptane/buffer (*35*) and the Triton X-100/n-hexanol/n-heptane/buffer (*35*), for further investigation (Figure 5-11). At the selected water-surfactant ratios w_0 of 10 and 15 for the AOT/n-heptane/buffer and the Triton X-100/n-hexanol/n-heptane/buffer systems, respectively, we found freezing points of –18°C for the AOT system and –26°C for the Triton system. We were able to take data up to 70°C in the Triton system, but only up to 30°C in the AOT system. Beyond 70°C in the Triton system, the reaction was too fast to measure ($T_{m,Triton}$ > 70°C). Beyond 30°C in the AOT system, data were too scattered to be interpretable (Figure 5-5). From this data we were able to determine the activation energy from the Arrhenius plot (Figure 5-6) as well as the activation parameters (Table 5-2) from the Eyring plot (Figure 5-7).

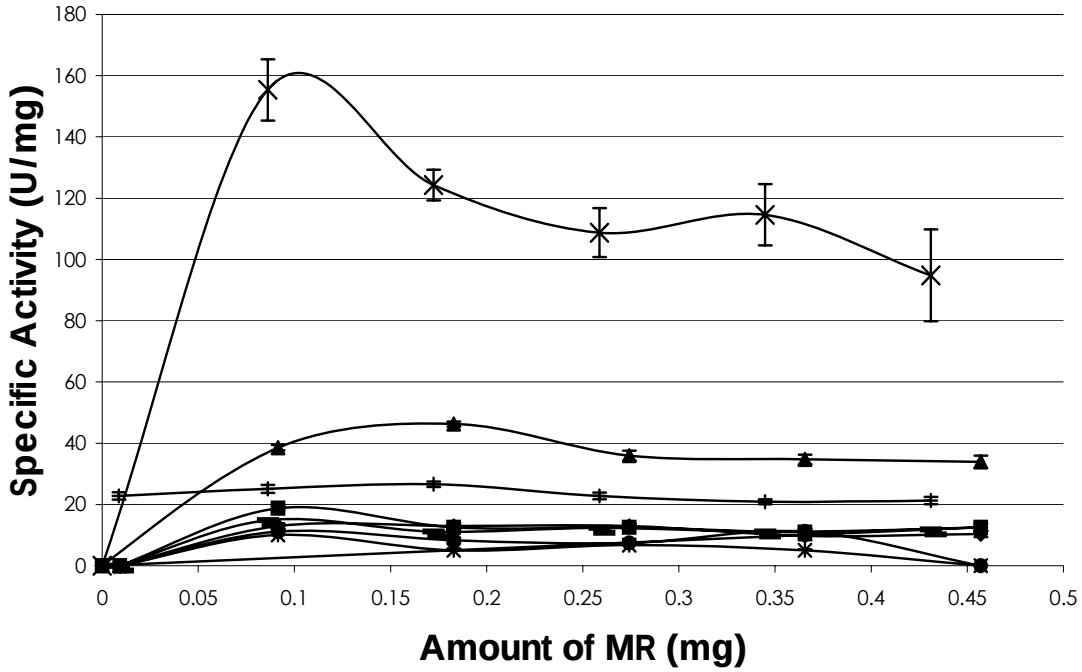


Figure 5-11. Specific activity of MR in various water-in-oil microemulsion micelle systems. (♦)Triton system, Tris buffer (pH 7.0); (■)Triton system, sodium phosphate (pH 7.5); (▲)Triton system, bistris propane (pH 6.5); (✕)Triton system, sodium bicarbonate (pH 7.0), (-) AOT system, tris (pH 7.5), (+) AOT system, phosphate (pH 7.0), (ж)AOT system, bistris propane (pH 6.5), (•) AOT system, sodium bicarbonate (pH 7.0).

5.4 Discussion

One constraint in the measurement of any enzyme activity at low temperatures is finding a liquid reaction medium below 0°C. The most common strategy to investigate enzyme activity at subzero temperatures is to employ organic cryosolvents, i.e. to change the medium by adding a water-soluble organic co-solvent, such as ethyleneglycol (EG), dimethylsulfoxide (DMSO), or ethanol. Our results in these media indicate

that MR is not stable in the presence of organic solvents, which affirms previous results (*31*).

Employing concentrated ammonium salt solution as a cryosolvent for low temperature studies was successful. We observed freezing points of 50% saturated ammonium acetate and ammonium formate at around -30ºC and found that MR retains activity at a similarly low temperature. We also observed a freezing point depression of nearly 5.2 ± 2.1 °C with the mandelate buffer (pH 7.5). At salt levels greater than 50% saturation, MR was inactive.

We observed in both the Arrhenius and Eyring plots (Figures 5-6 and 5-7, respectively) that the curves are not quite linear. Convex Arrhenius and Eyring plots suggest that there are two competing enzymatic forms (*36*), however it is unclear if that is the reason in this case. The activation parameters of MR racemization in concentrated salt solutions are somewhat similar to aqueous buffer (Table 5-2): activation energy E_a is within 3 kJ/mol and the enthalpy of activation $\Delta H^{\neq}$ is within 3-4 kJ/mol. The MR reaction in concentrated salt solutions seems to be slightly more entropically demanding than in water with a $\Delta S^{\neq}$ of –197 J/mol/K, where the $\Delta S^{\neq}$ in water is about –118 J/mol/K. Remarkably, the data for both ammonium salt solutions are almost identical, pointing to similar changes in enzyme reactivity in comparison with aqueous buffer. This effect suggests that the ammonium possibly has an effect, however it has been

shown that anion effects dominate (*28*), hence this could be explained by the chemical similarity between the acetate and formate anions.

The order of deactivation was first order, with the $k_{d,obs}$ ranging from 0.02 to 0.0001. The trend of the log of our experimentally measured deactivation constants (log $k_{d,obs}$) and B-viscosity coefficients are similar to other proteins such as horse liver alcohol dehydrogenase(HL-ADH), a-chymotrypsin, and monomeric red fluorescent protein (*28*). These results suggests that chaotropic solutions are much more deactivating than kosmotropic solutions. Therefore, the kosmotropic solutions such as ammonium acetate, ammonium sulfate, and ammonium formate are good ones for low temperature studies done in this work, especially since these salts seem to be somewhat stabilizing. As ammonium, acetate, formate, and sulfate are all present on plants and moons (*37*, *38*), showing enzymatic activity in environments is a promising sign that extraterrestrial life could exist.

We also discovered that MR is active in two microemulsion systems, the anionic surfactant Aerosol OT and the non-ionic surfactant Triton X-100 (20% n-hexanol as cosurfactant), both with n-heptane as organic continuum. The highest levels of MR activity were obtained at different water-surfactant ratios w_0 and in different buffers (Figure 5-11): while the maximum of the AOT system in phosphate buffer exhibited a shallow maximum at w_0 of 10, the maximum of the Triton X-100 system in

bicarbonate buffer at w_0 of 15 was very pronounced. In general, however, reactivity in microemulsions is decreased by a factor of five compared to water (*39*).

Comparing activation parameters of MR racemization in microemulsions with those in aqueous buffer (Table 5-2), we find while the entropy of activation $\Delta S^{\neq}$ is slightly decreased compared to aqueous buffer, the enthalpy of activation $\Delta H^{\neq}$ is somewhat close to that of the aqueous buffer, resulting in a similar value to buffer for the Gibbs free enthalpy of MR racemization $\Delta G^{\neq}$. The values for the Triton X system differ slightly more from aqueous buffer than those for AOT.

While we measured a temperature of optimum activity T_{opt} in aqueous buffer of 50° C, this value increased in both concentrated ammonium salt solutions to 60° C ± 5° C. In Triton X-100-containing microemulsions, T_{opt} was elevated even further, to beyond 70° C, at which point the reaction became too fast to follow with the endpoint assay employed in this work. [Data in the AOT system beyond 30° C were too scattered to be interpreted.] Thus, MR was stabilized with respect to temperature in both concentrated salt solutions and microemulsions compared to aqueous buffer.

5.5 Conclusion

The study enzyme reactivity at low temperatures requires a variety of changes from the methods used to measure enzyme reactivity in

ambient conditions. Since measurements of enzymatic rates at low temperatures cannot be performed in a purely aqueous environment, it is necessary to conduct the reaction in a medium with a melting point below the cold temperature environment in question. *Organic cryosolvents* are the most common medium to attempt these reactions at low temperature; however, such solvents tend to denature the protein or at least tend to vastly increase the K_M value, and therefore are not favorable to the enzymatic environment (*1*).

Another option is the use of highly concentrated aqueous salt solutions, such as saturated ammonium salts. Such a system has been used successfully to measure β-lactamase activity down to –60°C in saturated ammonium acetate solution (*2*). We were able to measure enzymatic activity down to -30°C in 50% saturated concentrated ammonium salts. We found similar activation parameters as those seen in the aqueous system, and as with the aqueous system, there was no observed mechanism change at 0°C. The kosmotropic salts seem to be stabilizing for the protein.

We also measured mandelate racemase activity in water-in-oil microemulsions in a temperature range between –30°C and 60-70°C. Thus, we have demonstrated both high concentrated salt solution and water-in-oil microemulsions, believed to be able to form on extraterrestrial

planets and moons in the presence of small amounts of liquid water, are suitable media for enzyme reactions at subzero temperatures.

5.6 REFERENCES

(1) Douzou, P. (1986) Interactive Effects of Cryosolvents, Ionic and Macromolecular Solutes on Protein Structures and Functions. *Cryobiology 23*, 38-47.

(2) Cartwright, S. J., and Waley, S. G. (1987) Cryoenzymology of Beta-Lactamases. *Biochemistry 26*, 5329-5337.

(3) Reat, V, Finney, J.L., Steet, A., Roberts, M.A., Smith, J., Dunn, R., Peterson, M., and Daniel, R. (2000) Cryosolvents Useful for Protein and Enzyme Studies Below -100°C. *J. Biochem. Biophys. Methods 42*, 97-103.

(4) Monnard, P. A., and Deamer, D. W. (2002) Membrane Self-Assembly Processes: Steps toward the First Cellular Life. *The Anatomical Record 268*, 196-207.

(5) Chang, G.G., Huang, T.M., and Hung, H.C. (2000) Reverse Micelles as Life-Mimicking Systems. *Proc. Natl. Sci. Counc. 24*, 89-100.

(6) Segre, D., Ben-Eli, D., Deamer, D. W., and Lancet, D. (2001) The Lipid World. *Origins of Life and Evolution of the Biosphere 31*, 119-145.

(7) Chang, T. (1964) Semipermeable Microcapsules. *Science 146*, 524-525.

(8) Shield, J. W., Ferguson, H. D., Bommarius, A. S., and Hatton, T. A. (1986) Enzymes in Reversed Micelles as Catalysts for Organic-Phase Synthesis Reactions. *Industrial & Engineering Chemistry Fundamentals 25*, 603-612.

(9) Orlich, B., and Schomacker, R. (2002) *Enzyme Catalysis in Reverse Micelles*, Vol. 75.

(10) Luisi, P. L., Giomini, M., Pileni, M. P., and Robinson, B. H. (1988) Reverse Micelles as Hosts for Proteins and Small Molecules. *Biochimica Et Biophysica Acta 947*, 209-246.

(11) Bommarius, A. S., and Hatton, T. A. (1995) Xanthine-Oxidase Reactivity in Reversed Micellar Systems - a Contribution to the Prediction of Enzymatic Activity in Organized Media. *J. Am. Chem. Soc. 117*, 4515-4523.

(12) Niemann, H. B., Atreya, S. K., Bauer, S. J., Carignan, G. R., Demick, J. E., Frost, R. L., Gautier, D., Haberman, J. A., Harpold, D. N., Hunten, D. M., Israel, G., Lunine, J. I., Kasprzak, W. T., Owen, T. C., Paulkovich, M., Raulin, F., Raaen, E., and Way, S. H. (2005) The Abundances of Constituents of Titan's Atmosphere from the Gcms Instrument on the Huygens Probe. *Nature 438*, 779-784.

(13) Mckay, C. P., and Smith, H. D. (2005) Possibilities for Methanogenic Life in Liquid Methane on the Surface of Titan. *Icarus 178*, 274-276.

(14) Douzou, P. (1977) *Cryobiochemistry: An Introduction*, Academic Press, New York, NY.

(15) Vajda, T. (1999) Cryo-Bioorganic Chemistry: Molecular Interactions at Low Temperature. *Cell. Mol. Life Sci. 56*, 398-414.

(16) Travers, F., and Barman, T. (1995) Cryoenzymology: How to Practice Kinetic and Structural Studies. *Biochimie 77*, 937-948.

(17) Bragger, J.M., Dunn, R.V., and Daniel, R.M. (2000) Enzyme Activity Down to -100°C. *Biochimica et Biophysica Acta 1480*, 278-282.

(18) King L. and Weber, G. (1986) Conformational Drift and Cryoinactivation of Lactate Dehydrogenase. *Biochemistry 25*, 3637-3640.

(19) Gerday, C., Aittaleb, M., Arpigny, J. L., Baise, E., Chessa, J. P., Garsoux, G., Petrescu, I., and Feller, G. (1997) Psychrophilic Enzymes:

A Thermodynamic Challenge. *Biochimica et Biophysica Acta 1342*, 119-131.

(20) Georlette, D., Blaise, V., Collins, T., D'amico, S., Gratia, E., Hoyoux, A., Marx, J. C., Sonan, G., Feller, G., and Gerday, C. (2004) Some Like It Cold: Biocatalysis at Low Temperatures. *Fems Microbiology Reviews 28*, 25-42.

(21) Pogorevc, M., Stecher, H., and Faber K. (2002) A Caveat for the Use of Log P Values for the Assessment of the Biocompatibility of Organic Solvents. *Biotechnology Letters 24*, 857-860.

(22) Braco, L. (1995) Biocatalysis and Biorecognition in Nonaqueous Media. Some Perspectives in Analytical Biochemistry. *120 1-4*.

(23) Russell, A. J., Chatterjee, S., and Bambot, S. (1992) Mechanistic Enzymology in Non-Aqueous Media. *Pure and Appl. Chem. 64*, 1157-1163.

(24) Walde, P. (2006) Surfactant Assemblies and Their Various Possible Roles for the Origins of Life. *Origins of Life and Evolution of the Biosphere 36*, 109-150.

(25) Torres, S., and Castro, G.R., (2004) Non-Aqueous Biocatalysis in Homogeneous Solvent Systems. *Food Technol. Biotechnol. 42*, 271-277.

(26) Simoupoulos, T.T., Jencks, W.P. (1994) Alkaline Phosphatase Is an Almost Perfect Enzyme. *Biochemistry 33*, 10375-10380.

(27) Laane, C., Boeren, S., Vos, K., and Veeger, C. (1987) Rules for Optikmization of Biocatalysts in Organic Solvents. *Biotechnol. Bioeng. 30*, 81-87.

(28) Broering, J.M., and Bommarius, A.S. (2005) Evaluation of Hofmeister Effects on the Kinetic Stability of Proteins. *J. Phys. Chem. B 109*, 20612-20619.

(29) Jones, G., and Dole, M. (1929) *J. Am. Chem. Soc. 51*, 2950.

(30) Castro, G. R., and Knubovets, T. (2003) Homogenous Biocatalysis in Organic Solvents and Water-Organic Mixtures. *Crit. Rev. in Biotech. 23*, 195-231.

(31) Kaftzik, N., Kroutil, W., Faber, K., and Kragl, U. (2004) Mandelate Racemase Activity in Ionic Liquids: Scopes and Limitations. *Journal of Molecular Catalysis A: Chemical 214*, 107-112.

(32) Buzko V.Y., Poluskin, A. A., and Sukhno I.V. (2004) in *Equilibrium Reaction Constants in Aqueous Solutions* (V.Y., P. L. D. S. I. V. B., Ed.), IUPAC.

(33) Kaftzik, N., Kroutil, W., Faber, K., and Kragl, U. (2004) Mandelate Racemase Activity in Ionic Liquids: Scopes and Limitations. *Journal of Molecular Catalysis A: Chemical 214*, 107-112.

(34) Pogorevc, M., Stecher, H., and Faber, K. (2002) A Caveat for the Use of Log P Values for the Assessment of the Biocompatibility of Organic Solvents. *Biotechnology Letters 24*, 857-860.

(35) Bommarius, A. S., and Hatton, T. A. (1995) Xanthine-Oxidase Reactivity in Reversed Micellar Systems - a Contribution to the Prediction of Enzymatic Activity in Organized Media. *J. Am. Chem. Soc. 117*, 4515-4523.

(36) Truhlar, D. G., and Kohen, A. (2001) Convex Arrhenius Plots and Their Interpretation. *Proc Natl Acad Sci U S A 98*, 848-851.

(37) Zolotov, M. Y., and Shock, E. L. (2004) A Model for Low-Temperature Biogeochemistry of Sulfur, Carbon, and Iron on Europa. *Journal of Geophyscial Research 109*, E06003-06001-06016.

(38) Spaun, F., Schmidt, J., Albers, N., Horning, M., Makuch, M., Seib, M., Kempf, S., Srama, R., Dikarev, V., Helfert, S., Moragas-Klostermeyer, G., Krivov, A. V., Sremcevic, M., Tuzzolino, A. J., Economou, T., and

Grun, E. (2006) Cassini Dust Measurements at Enceladus and Implications for the Origin of the E Ring. *Science 311*, 1416-1418.

(39) Douzou, P., Keh, E., and Balny, C. (1979) Cryoenzymology in Aqueous Media: Micellar Solubilized Water Clusters. *PRoc. Natl. Acad. Sci. 76*, 681-684.

CHAPTER 6

Evolution of Enzymatic Activity in the Enolase Superfamily: three mutations create cross-reactivity between two subgroups

6.1 Introduction

Upon origination, life most likely began to diversify through natural selection, or evolution. When investigating the existence of extraterrestrial life, it is of interest to understand how this evolution occurred and/or continues to occur. The concept of evolution was first introduced in "The Origin of Species" by Charles Darwin (*1*). He hypothesized that complexity evolves through an elaborate and optimized stepwise process of natural selection. This however does not explain the evolution of a lock-and-key system (such as enzymes and hormone receptors), as evolution of a lock-and-key system would require two simultaneous changes (the enzyme and substrate). Darwin was aware that if these types of systems existed in biology, his theory would lose merit, he remarked "if it could be demonstrated that any complex organ existed which could not possibly have been formed by numerous successive slight modifications, my theory would absolutely break down (*1*)." Darwin also suggested these issues could be resolved "if we look to an organ common to all members of a

large class...in order to discover the early transitional grades through which the organ has passed, we should have to look to very ancient ancestral forms, long since become extinct" (*1*). Therefore, the path of evolution will be masked if only extant forms are observed; in turn this leads to an incomplete picture of the history of a protein. Herein lies a problem in that there is an absence of "protein fossils," which prevents a clear picture of protein evolution to be observed (*2*).

In 1970, Ohno developed a model that suggests enzymes have evolved through gene duplication (*3*). A few years later, Jensen proposed that enzymatic promiscuous functions provides a template for the evolution of new enzymes (*4*). Enzymatic evolution through promiscuity has been documented (*5-7*), Khersonsky et. al. have developed a model showing that weak negative trade-offs allow for divergence of a new function through a "generalist" intermediate, that has a broader specificity (*5*). Therefore, through the detection of promiscuous enzymes, or cross-reactivities between family members one could reveal evolutionary relationships within the enzyme families (*5*).

A computational study was done using digital organisms to construct the complete evolutionary history of a complex function (*8*). The group found that Darwin's hypothesis is valid and complex features evolve through the modification of existing structures and functions (*8*). Another group has demonstrated the evolution of a steroid hormone-

receptor pair through a stepwise process, in which they were able to assemble various integrated systems through combinations of the old molecules with new ones having slight modifications from the old ones (*2*). This concept can also be applied to find ancestral enzymes that will provide insight into the evolution of enzyme superfamilies. In order to prove that such systems gradually evolve, we chose the enolase superfamily as a model system.

The enolase superfamily has been used as a model system to describe enzyme evolution in many cases as it has been extensively characterized and is extremely mechanistically diverse, catalyzing at least fourteen different reactions (*9-14*). Members of this diverse superfamily are classified based on the capability to catalyze the initial abstraction of the α-proton of the carboxylate substrate to generate and stabilize an enolate anion intermediate (*15*). Although members of this superfamily catalyze very different types of reactions, nearly all of them share three conserved residues for binding the necessary divalent metal ion (*16*). The overall sequence identity of the members in this superfamily is low (generally less than 25% in pair-wise alignments of the barrel domains) although structural characteristics are conserved (*11*). Enolase superfamily members share the same bidomain structure, where the catalytic domain is a modified $(\beta/\alpha)_8$-barrel with conserved catalytic residues located at the C-terminal ends of the β-strands, and the second

domain acts as a cap on the active site determining the shape and polarity of the active site cavity (*10, 13*).

As reconstruction of the evolutionary pathway of proteins is complicated due to a lack of "protein fossils;" growing evidence has accumulated to support that natural divergence of enzyme families and superfamilies has proceeded through highly promiscuous, progenitors. In some cases, ancient genes have been resurrected to aid in reconstructing evolution of lock-and-key proteins (*2*).

In this thesis we provide the first evidence of possible divergent evolution across the MLE and MR subgroups of the enolase superfamily; so far, evolution has been demonstrated within subgroups. We found that MR shows very slight activity on phenylglycine and hypothesized, as the tertiary structures of MR and MLE are superimposable within 1 Å, especially in the $(\beta/\alpha)_7\beta$-barrel domain, that we also can observe N-acylamino acid racemase (NAAR) activity in MR or its variants and thus demonstrate cross-reactivity across the MR and MLE subgroups. A potential divergent evolution pathway consists of two parts: i) obtaining the new function in MR, and ii) exchanging the active site residues between MR and MLE. In this work, we deconvolute both processes and first create a variant of MR that can catalyze a reaction normally performed only by members of the MLE subgroup (racemization of an N-acylamino acid). Secondly, we demonstrate that the active site residues of the MLE subgroup (two Lys)

can be introduced into the above-described MR-variant (active site Lys and a His-Asp dyad, see Figure 1), without dramatic loss of activity. These experimental results support the hypothesis that different enolase superfamily subgroups may have evolved by divergent evolution.

6.2 Background

6.2.1 Evolution Models

Although enzymes are robust and highly proficient, they are highly evolvable in order to adapt to changing environments. Based on Darwin's thoughts that evolution occurs gradually, Jensen formulated a hypothesis in 1974, stating that enzymes exhibiting broad specificity provided starting points for modern enzymes (*4*). This suggests that ancestral enzymes possessed broad specificity with the ability to perform a multitude of functions necessary to maintain ancestral organisms. Specialized enzymes that increased metabolic efficiency evolved through gene duplication and divergence. Once a latent promiscuous function becomes relevant, it can be improved through one, or just a few mutations, to provide a distinct selection advantage. In most cases, the goal of directed evolution experiments is to further evolve promiscuous activity, typically a substrate or reaction, that bares a significant resemblance to the original function. The main conclusion from these experiments was that these promiscuous functions exhibit a high

"plasticity" where a few mutations can readily increase a promiscuous activity, typically by 10-10000 fold (*5*). Another interesting aspect of promiscuous functions is that, in many cases, improvements in promiscuous functions do not seem to trade-off with a parallel decrease in the original function (*17*). Eighteen cases (*17*) and a further eleven cases (*5*) showed an increase in promiscuous activity with 1-11 mutations under selection by 10-10^6 fold, whereas the original activity of these protein decrease by only 0.8-11 fold. The native function was conserved despite the application of only one selection criterion. This observation has important implications regarding the understanding of the early steps in the evolution of protein functions.

There is no doubt that gene duplication is a necessary step but the widely accepted Ohno model (*3*), surmises that the generation of a redundant gene copy that is relieved from the burden of selection is the first step in the evolution of a new function. The gene sharing model is another model that describes the divergence of new function proceeds via a 'generalist' intermediate that exhibits broad specificity. This process can then be followed by gene duplication leading to divergence of a 'specialist.' Most of the directed evolution experiments described above agree with the gene-sharing model but there are also examples that agree with Ohno's model. These are found if a dual selection pressure is applied, such as a parallel selection for increase of a promiscuous activity

with the application of a decrease in the native activity (*18, 19*). Based on the gene sharing model and Ohno's model, Tawfik described a model where the relative rates matter by which a new function is gained, and the old function is lost (*5*) as shown in Figure 6-1. This model makes the assumption that trade-offs can be determined and quantified, in particular with enzymatic activities. Therefore, in those cases where the negative trade-off is weak, the divergence of new function can proceed via a 'generalist' intermediate that exhibits broad specificity. Gene duplication can then follow this process, rather than initiating it, leading to divergence of a new 'specialist' (*5*). As previously stated, based on the lack of "protein fossils" the reconstruction of protein evolution pathways is difficult. However, hints regarding evolutionary and mechanistic relationships within enzyme families can be provided through identifying promiscuous activities, or cross-reactivities, between different members of the same enzyme family or superfamily (Figure 6-2).

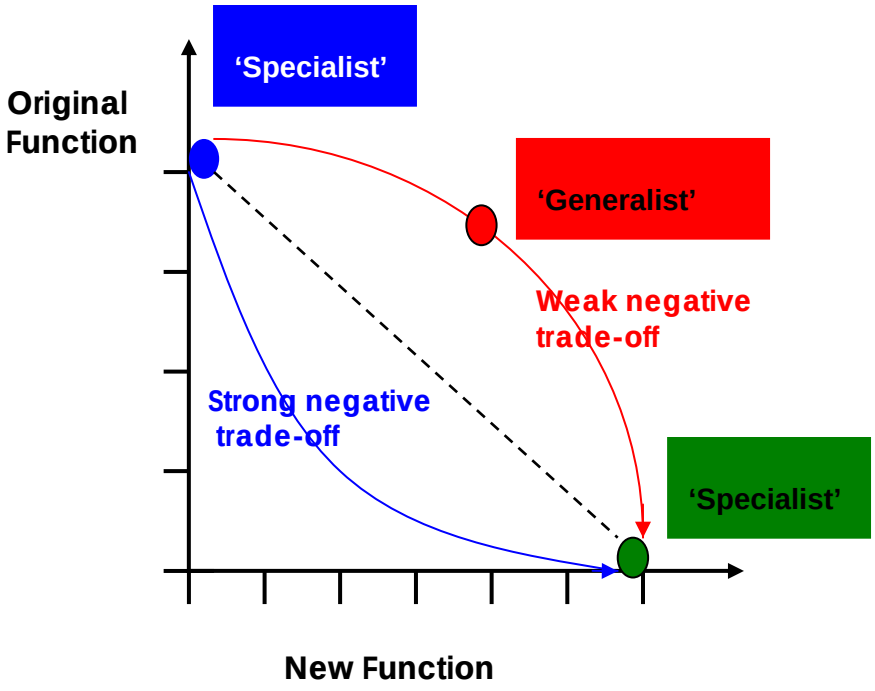


Figure 6-1. Trade-off's involved in conversion of one 'specialist.' Figure adapted from Khersonsky et. al. (2006).

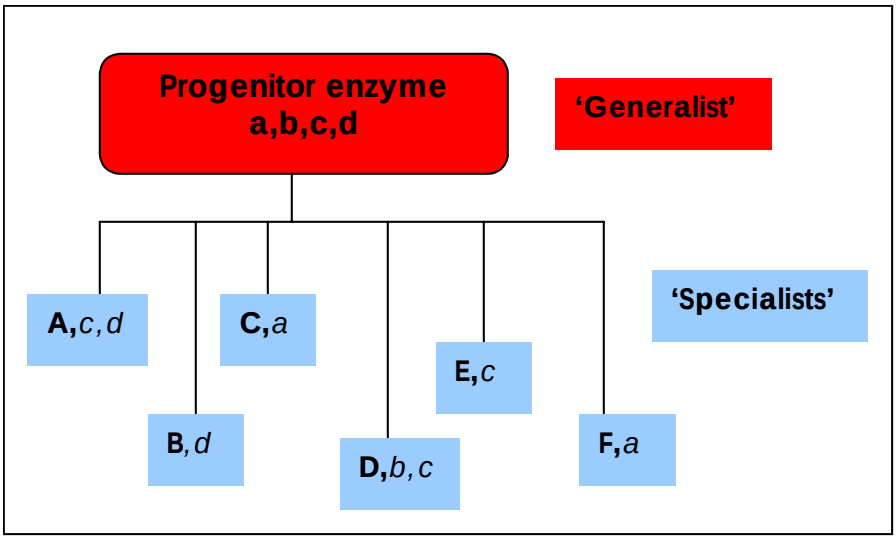


Figure 6-2. Divergence of new enzyme family members in nature. Figure adapted from Khersonsky et. al. (2006).

6.2.2. Enolase Superfamily

Proteins containing the α/β-barrel structure are the most abundant group with 10% of all enzymes with known structure, providing a scaffold for diverse enzymatic functions (*20, 21*). The active site in α/β-barrel enzymes structure tends to be at the C-terminal ends of the β-strand, hence it has been suggested that enzymes of this fold evolved from a small number of progenitors (*21, 22*). Structural residues of this type are separate from residues responsible for catalysis, an advantage when evolving an enzyme with a new function while keeping the structure unchanged. New functions have been acquired by single residue mutations in α/β-barrel enzymes, rendering enzymes with the α/β-barrel structure desirable for investigation of the evolutionary potential of enzymes (*23*).

The enolase superfamily is classified as a cousin of the α/β-barrel proteins, because the structure contains the α/β-barrel domain but unlike the α/β-barrel protein family the enolase superfamily has an additional capping domain (Figure 6-3). Enzymes in the enolase superfamily catalyze at least fourteen different chemical reactions, including racemization, epimerization, and both *syn* and *anti* β-elimination reactions involving water, ammonia, or an intramolecular carboxylate group as a leaving group (*15*). Based on primary sequence alignments, members of the enolase superfamily are assigned to four subgroups: the MR

subgroup, the MLE subgroup, 3-methylaspartate ammonia lyase (MAL), and the enolase subgroup. The four subgroups can then be identified by: the identities and sequence contexts of the three carboxylate ligands for the essential divalent metal ligand; the number of identities of the general base catalysts; and identities of the electrophilic catalysts that stabilize the enolate anion intermediate (*24*).

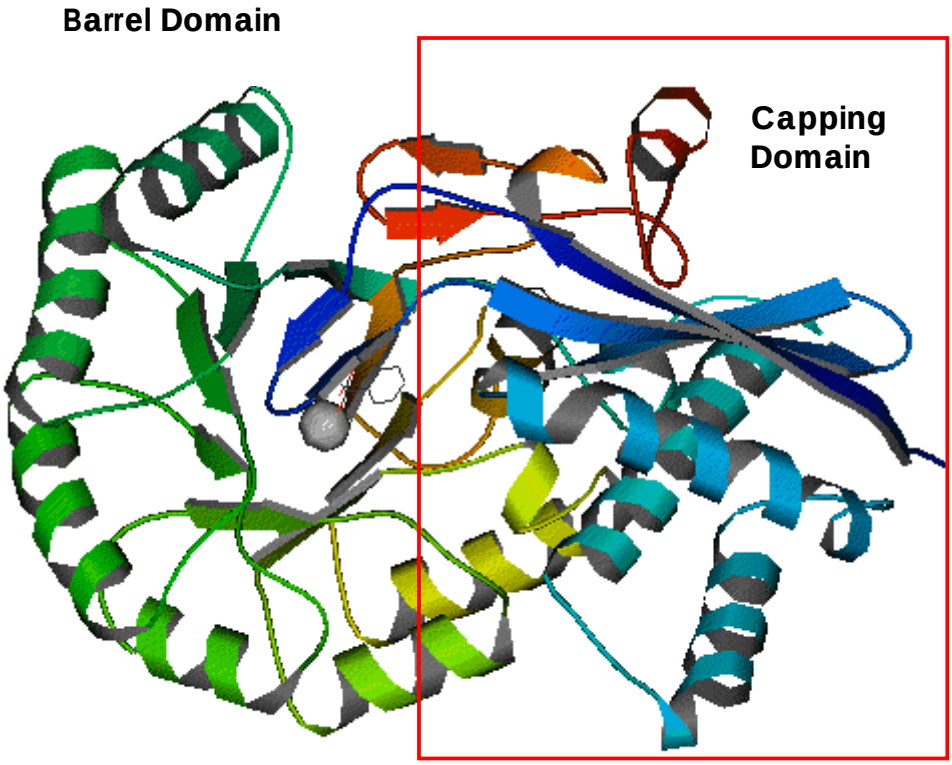


Figure 6-3. Tertiary structure of MR (1MNS), showing capping and barrel domains.

The structures of members of the enolase family are known to be nearly superimposable. For example, MLE and MR, are known to be nearly superimposable, with a root-mean-square distance of 1.3 Å between equivalent Cα atoms on the tertiary structure level, although they have low identities of about 26% at the amino acid level (*25*).

6.2.1.1 Divergent evolution in the enolase superfamily

Members of the enolase superfamily can be grouped into three distinct active site contexts, based on the identities and positions of shared general acid/base catalysts in the $(\beta/\alpha)_8$-barrel domain: 1) both the enolase and MAL subgroup in which a Lysine is located at the C-terminal and of the sixth β-strand; the muconate lactonizing enzyme (MLE) subgroup in which the Lysine residues are located at the C-terminal ends of both the second and the sixth β-strand; and the mandelate racemase (*26*) subgroup in which a Histidine-Aspartate dyad is located at the C-terminal end of the seventh and sixth β-strand and a lysine at the C-terminal end of the second β-strand, respectively, see Figure 6-4.

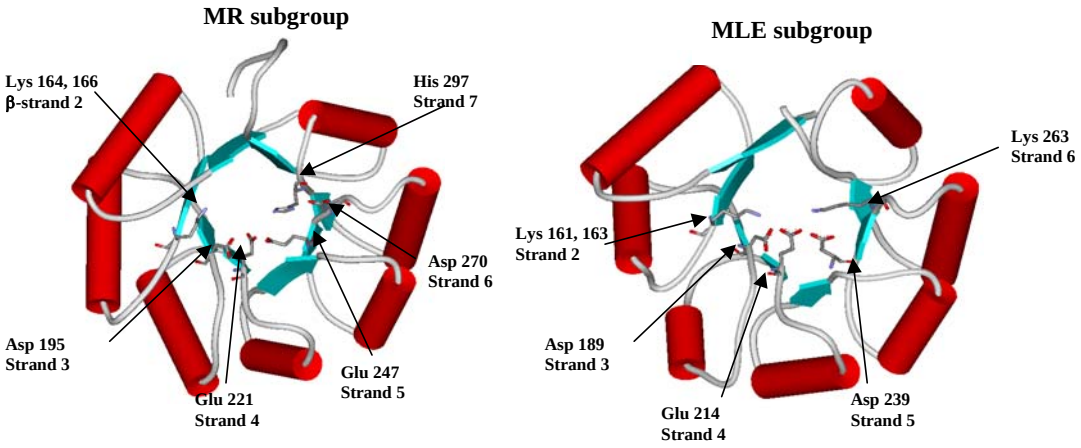


Figure 6-4. Comparison of active site residues from two enolase superfamily subgroups. (A) mandelate racemase (MR) from PDB file 1MDR and (B) *o*-succinylbenzoate synthase (OSBS/NAAR) from PDB file 1SJD. Active site residues of MR are mutated to the structural homologues of OSBS: D270K; H297G.

Convincing evidence for the divergent evolution from a common progenitor in the enolase superfamily is provided by the conservation of the bi-domain structure (*11*). The extraordinarily divergent capping domain on the other hand complicates the identification of evolutionary origins, as shown for the OSBS/NAAAR family. No sequence or structural motifs unique to this family could be identified (*27*). The functional assignment of new genes is difficult for this family, since the same chemical reaction with different substrate specificities, can be catalyzed by members of different subgroups (*12*). Additionally, some genes have shown promiscuity and therefore mislead the assignment of the native function (*28*). Further evidence for divergent evolution has been recently provided in three cases by the Gerlt group: 1) The *o*-succinylbenzoate synthase (OSBS) from *Amycolatopsis* (2.5×10^5 M^{-1} s^{-1}) (*28*) also shows N-acylamino acid racemase activity (NAAR) (370 M^{-1} s^{-1}). Even though this later activity is considered to be accidental catalytic promiscuity (*12*), it demonstrates the evolutionary potential of this enzyme class. The structure-based explanation is that both substrates bind in a hydrophobic pocket, formed by residues of the capping domain (*29*). 2) Further proof of the evolutionary potential was given by the identification of a single amino acid substitution in both the L-Ala-D/L-Glu epimerase from *Escherichia coli* and muconate lactonizing enzyme II (MLE II) from *Pseudomonas putida* that led both enzymes to catalyze the *o*-

succinylbenzoate reaction, as well as their own starting reaction, see Figure 6-5 (*30*). This demonstrated that the activities of the MLE subgroup could be converted in some cases by as few as one mutation. 3) Recently, another promiscuous activity has been observed in the MLE subgroup. It has been shown that an enzyme from *Geobacillus kaustophilus* was an efficient N-succinylamino acid racemase (NSAR), and *o*-succinylbenzoate synthase (OSBS), respectively (NSAR: up to 3.7×10^4 M^{-1} s^{-1}; OSBS: 1.9×10^6 M^{-1} s^{-1}) (*12*). Unexpectedly, the enzyme was not encoded in the OSBS-typical menaquinone biosynthesis operon, if not in a new operon. Based on the gene environment and the characterization of additional enzymes, they hypothesized that this operon encodes an irreversible pathway for the conversion of D- to L-amino acids. Further they hypothesized that this could be a bifunctional enzyme rather than an accidental promiscuous enzyme, based on the absence of an OSBS gene in the OSBS-typical operon and the efficient cataysis of the OSBS reaction by the NSAR (*12*).

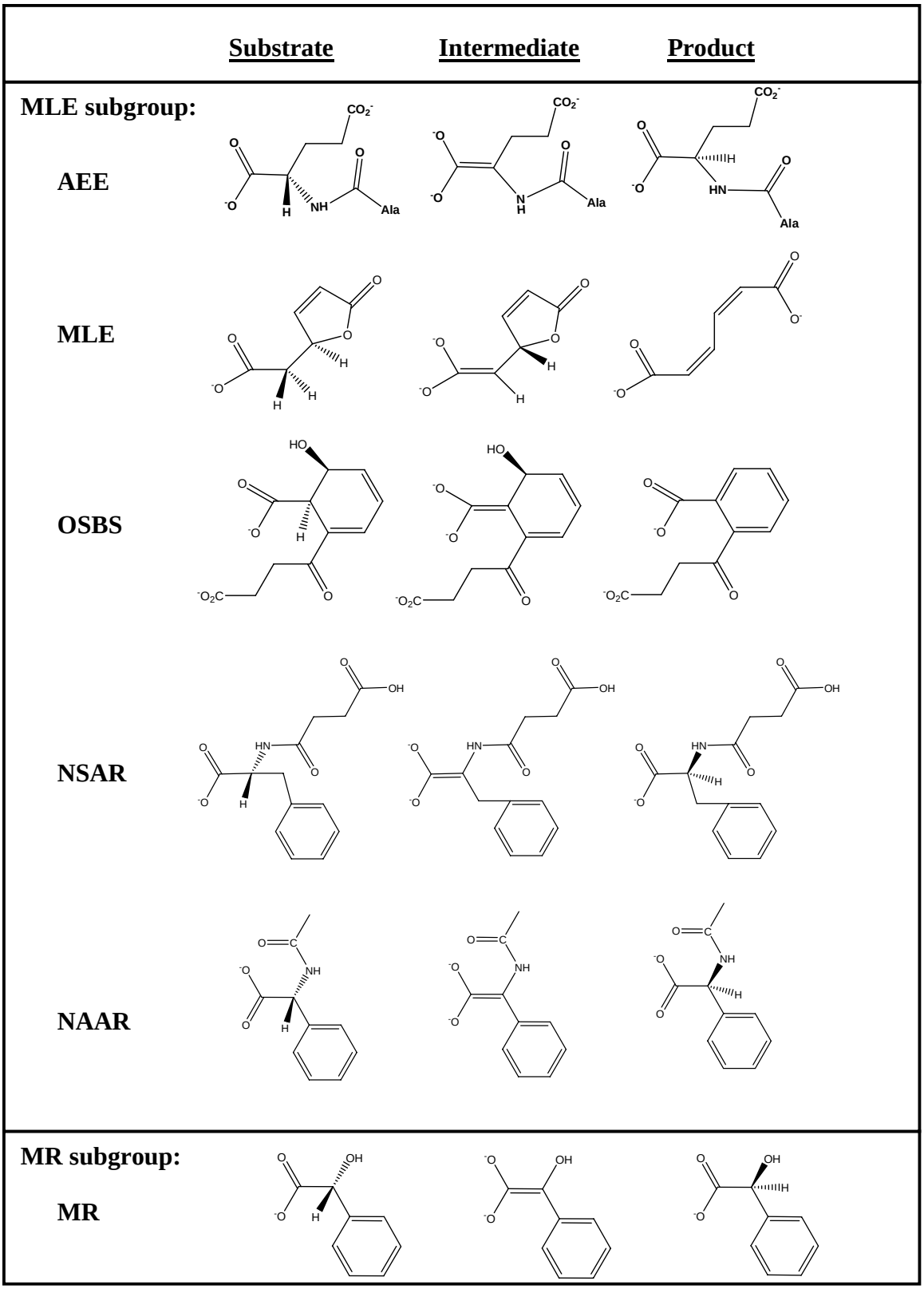


Figure 6-5. Reactions catalyzed by the MR and MLE subgroups.

6.3 Materials and Methods

6.3.1 Materials

(*R*)-mandelic acid was a kind gift of BASF (Ludwigshafen Germany); (*S*)-mandelic acid, cis,cis muconate, *S*-methionine, N-acetyl-(*R*)-methionine, N-acetyl-(*S*)-methionine, (*R*)-phenylglycine, and Tris(hydroxymethyl)-aminomethane were obtained from Sigma-Aldrich (St. Louis, Missouri), $MgCl_2$ was from Fisher Chemical (Fair Lawn, New Jersey), Luria Broth, lysozyme, are from EM Science (Gibbstown, New Jersey), ampicillin from Shelton Scientific (Shelton, Connecticut), L-phenylglycine from Alfa Aesar (Ward Hill, Massachusetts), N-formyl-phenylglycine, and N-methyl-(*R*)-phenylglycine, N-formyl-(*R*)-phenylglycine, and N-formyl-(*S*)-phenylglycine from Bachem (King of Prussia, Pennsylvania).

6.3.2 Bacterial Strains and Growth Conditions

The bacterial strains used in the study are located in Table 6-1. The *E. coli* strains, which were used for proliferation of plasmids, were grown at 37°C in Luria-Bertani (LB) media (*31*) or Terrific Broth (*32*), containing either 100 μg/ml ampicillin or 30 μg/ml kanamycin depending on the plasmid. Solid medium was made up of LB medium containing 1.5% of agar.

Table 6-1. Bacterial Strains used in this study.

Strain	Characteristics	Source
E. coli XL1-Blue	Gene cloning host: *recA1 endA1 gyrA96 thi-1 hsdR17 supE44 relA1 lac [F' proAB lacI^{q}ZΔM 15 Tn10 (Tet')]*	Stratagene
E. coli JM105	Gene cloning host: *endA*1, *thi, rps*L, *sbcB*15, *hsdR*4, D(*lac-proAB*), [F8, *traD*36, *proAB*, *lacI*q *Z*DM15]	Stratagene

6.3.3 General DNA Techniques

Extraction of plasmid DNA was prepared using the QIAprep Spin Miniprep Kit (Qiagen, Valencia, CA). DNA was cleaned using the DNA Clean & Concentrator kit (Zymo Research, Orange, CA). Agarose gel electrophoresis was setup according to Sambrook (2001) (*31*) using a 1.2% TAE agarose gel. Agarose gels were stained with GelStar® Nucleic Acid Gel Stain (BMA, Rockland, ME) according to manufacturers instructions. Fragments of DNA were isolated from the agarose gel using the QIAquick Gel Extraction Kit or Qiaex II (Qiagen, Valencia, CA) as well as 1 kb ladder (New England Biolabs, Beverly, MA) was used for molecular weight standards.

Restriction enzymes were used according to the suppliers *Nde*I, *Bam*HI (New England Biolabs, Beverly, MA), *Eco*RI, *Hind*III, and *Dpn*I (Fermentas, Hanover, MD). Ligations were done using two different kits and used according to the suppliers' instructions: DNA Ligation Kit

(Novagen, Madison, WI) and Quick Ligation Kit (New England Biolabs, Beverly, MA). Competent *E. coli.* XL1-Blue cells were prepared according the manufacturer; all plasmids were transformed according to cell manufacturer.

6.3.4 Overlap Extension Method

Mutants were constructed using the overlap extension method. This method involves two PCRs using external primers and the mutation primers. Both of the products are gel-purified and a secondary PCR with only external primers is used to combine the pieces. The pieces were then cut with *Eco*RI and *Hind*III and ligated to the vector pkk223 (same vector as the wild-type). The transformants were then analyzed for gene and sent for sequencing.

The PCR reactions were carried out in a programmable Mastercycler® gradient thermocycler (Eppendorf, Westbury, NY). Amplification was always done using a hotstart PCR program, the DNA was pre-denatured at 98°C for 5 minutes before adding the *Pfu* enzyme (used according to manufacturer). The following sequence was used for 25 cycles: 95°C for 30 seconds, 5°C less than T_M for 30 seconds, and 72°C for 1 minute 15 seconds. An elongation cycle was added at the end at 72 °C for 10 minutes. All PCR runs were conducted with this method unless noted otherwise.

Table 6-2. Primers used in this study. Underline=mutation, italic=restriction enzyme site.

	Forward primer (5')	Reverse primer (3')
Wild type	gcgc*gaattc*atgagtgaa gtactgattaccggcctg	gcgc*aagctt*ttacaccag atatttcccgatttcttt
H297G	cggtattccaatgtccagc ggcctgttccaagaaatcag c	gctgatttcttggaacagccg gctggacattggaataccg
H297K	cggtattccaatgtccagcaa actgttccaagaaatcagc	gctgatttcttggaacagttt gctggacattggaataccg
D270K	gcatgccggttggctatgcc aaaggcaatgaagatcggtg gc	gccaccgatcttcattgcctt t ggcatagccaaccggcatgc
E247D	gctcaatgtgcccgtccagat gg gtgataactggctcggccctg agg	cctcagggccgagccagtta tcacccatctggacgggcacat tgagc
F52V	cgtggtaggccattcctacct g gtggcatacacccccgttgc	gcaacgggggtgtatgccac caggtaggaatggcctaccac g
E317D/ L319A	ccaactgcgcattggctg gatcgtgcggatctcgccgg c	gccggcgagatccgcacgatc cagccaatgcgcagttgg

6.3.5 Synthesis of N-acetyl-phenylglycine and N-succinyl-phenylglycine

The synthesis for N-succinyl-phenylglycine was performed according to Sakai et. al., described in the supplementary information (*12*). N-acetyl-phenylglycine was synthesized using this same method with the substitution of the succinic anhydride with acetic anhydride (*12*).

6.3.6 Superimposition of NAAR on MR.

We overlayed the crystal structure of OSBS/NAAR from *Amycolatopsis* co-crystallized with N-succinyl phenylglycine (PDB: 1SJD) structure onto the crystal structure of MR co-crystallized with mandelate

(PDB: 1MDR) using Swiss-PDB Viewer, version 3.7. Four structurally and sequentially conserved amino acid residues were chosen to act as pivot points. These residues were determined based on their role in the active site of each gene, as well as the conservation of these residues in many of the enolase family members. The amino acid residues chosen were Asp (D195 in 1MDR and D189 in 1SJD); Glu (E221 in 1MDR and E214 in 1SJD); Lys (K166 in MDR and K163 in 1SJD); and Lys (K164 in 1MDR and K161 in 1SJD). The focus of the overlay was the active site, the conserved metal atoms from each structure were only 0.5 Å apart.

6.3.7 Polarimetric Assay for MR activity

All polarimetric measurements were taken on a Jasco P-1010 (Jasco Corp., Easton Maryland) at 390 nm using a 10 cm pathlength and a total volume of 8 mL. The reading of optical rotation versus time was recorded. Data was recorded at 24°C, 100 mM mandelate in an aqueous solution of 50 mM Tris buffer, pH 7.5, and 1 mM $MgCl_2$ to ensure adequate cofactor presence.

6.3.8 Polarimetric Assay for Racemase Activity with Other Substrates

The reading of optical rotation versus time was recorded. Data was recorded at 24°C, 16 mM substrate solution in an aqueous solution of 50 mM Tris buffer, pH 7.5, and 1 mM $MgCl_2$ to ensure adequate cofactor presence. All substrate solutions were assayed with each mutant for activity over a 10-minute period at 23°C, pH 7.5. If there was no activity in

this time frame the solutions were incubated for 24 hours at 30°C. Initial optical rotation measurements were taken for the substrate solutions. The following controls were run: substrate solution alone, lysozyme added at the same molar concentration as MR variant, heat denatured (75°C) MR variants were also incubated to make sure any phenomenon seen would be due to the mutant protein.

6.4 Results and Discussion

6.4.1 Mandelate racemase shows weak activity on phenylglycine.

Felfer et. al. (2005) (33) recently published results stating that MR is totally inactive on both a-phenylglycine and N-acetyl phenylglycine. They also cited that the reason could be lack of sensitivity of the polarimetric assay. In contrast, we found that MR has a very low activity at pH 8.0 and 25oC on both enantiomers of phenylglycine with 82% conversion in a 24-hour period. The clearly discernible change in optical rotation over time corresponds to a specific activity of 1.2 x 10-3 ± 2.4 x 10-4 U/mg for both enantiomers, after subtraction of a slight background chemical reaction of 5 x 10-4 U/mg +/- 2 x 10-4 U/mg. We also tested N-acetyl-phenylglycine as a substrate for MR but found no activity, as did Felfer et. al. (2005) (33).

Our starting template was MR (P. putida) from the MR subgroup and the target template NAAR (Amycolatopsis sp.) from the MLE subgroup. The question arose how N-acylamino acid racemase is able to stabilize

the amino group of amino acids while mandelate racemase apparently does not stabilize this group. To determine the orientation of the substrate in the binding pocket in its relation to the divalent metal ion, we chose to overlay the crystal structure of OSBS/NAAR from Amycolatopsis sp. (PDB: 1SJD) onto the crystal structure of MR from P. putida (PDB: 1MDR).

6.4.2 Substrates to investigate

Although the substrate specificity of MR allows variation in the phenyl ring part of mandelate, the α-hydroxy group is deemed essential for catalysis because it acts as an electron-withdrawing group (*33*), accentuating the positive electrostatic environment of the active site through interactions with the divalent metal ion. However, an amino group or even an acyl-substituted amino group replaces the hydroxyl group with a less electron-withdrawing one. If the partially positive character of the chiral α-carbon atom, besides steric demands, is decisive for MR catalysis substrates with gradually less partially positive character of the chiral α-carbon atom and increasing steric demands can be tested. Therefore, we chose to test both enantiomers of α-phenylglycine, both enantiomers of N-formyl-phenylglycine, N-methyl-(*R*)-phenylglycine, N-acetyl-(*R*)-phenylglycine, N-succinyl-(*R*)-phenylglycine, and N-acetyl-(*R*)-phenylalanine, and N-acetyl-(*R*)-methionine (Figure 6-6). These substrates were tested on wild-type MR as well as on the variants with increased size of active site cavity: besides both enantiomers of mandelate, only the two

enantiomers of α-phenylglycine were accepted by wild-type MR (Table 6-3).

6.4.3 Mutations to increase the size of MR active site.

Through the overlay of the two crystal structures, we identified amino acid residues responsible for structural clashes within the MR active site with N-acetyl-phenylglycine (shortened from the N-succinyl-phenylglycine used in 1SJD), which is not a substrate for MR (Figure 6-7). We also observed that N-acetyl-phenylglycine is positioned within the MR active site such that the divalent metal ion binds in a different orientation than to mandelate (Figure 6-8). The altered orientation allows the substrate to be stabilized without the involvement of the α-hydroxyl group in binding. We identified three amino acid residues, F52, E317, and L319, that have clashes with N-acetyl-phenylglycine, and, when mutated, provide a larger binding pocket for N-acetyl-phenylglycine in MR (Figure 6-8). Both F52 and E317 were mutated to the structurally homologous amino acid in NAAR (F52V and E317D). E317 as the general acid catalytic residue in MR is positioned for partial proton transfer to the charge-neutralized carboxylate group of mandelate, forming a low barrier hydrogen bond stabilizing the enolic intermediate (34). As there are significant differences in the structure of the capping domains between MR and NAAR, there was no structural homolog for L319 (there is a large difference in the backbone between the two proteins at this point).

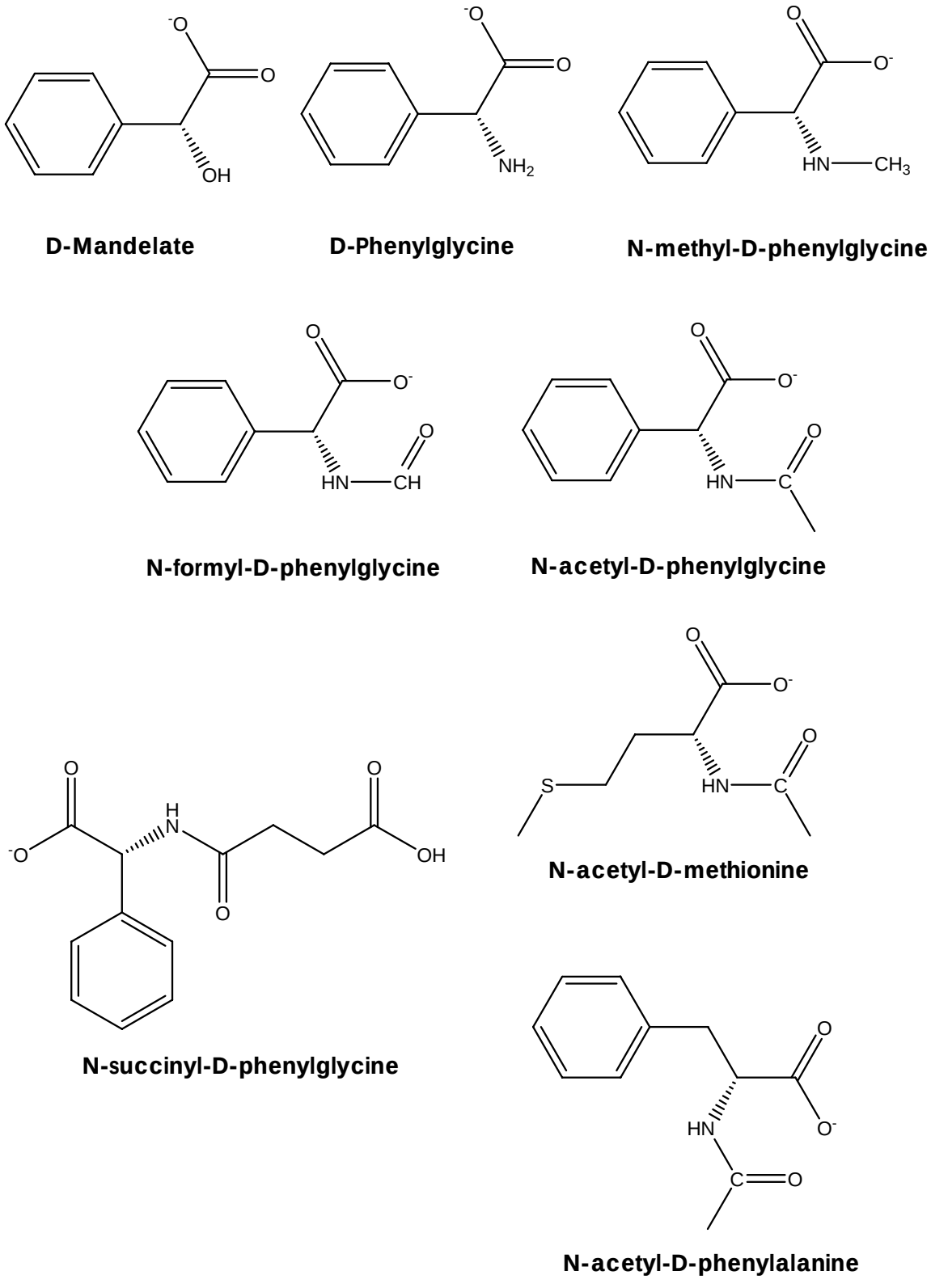


Figure 6-6. Substrate series to test to provide evidence for divergent evolution in the enolase superfamily.

Hence, we could not imitate the NAAR structure simply through a point mutation, so we chose to mutate this residue to a small amino acid (L319A). The F52V mutation is important because it is located in the hydrophobic (*S*)-pocket, interacting with the phenyl group, therefore, the substitution should be to another hydrophobic amino acid. It has been shown that mutations of this and another residue to a bulkier amino acid, F52W/F54W, led to reduced greatly reduced k_{cat} for both (*R*)- to (*S*)- and (*S*)- to (*R*)- directions (*35*).

6.4.3.1 Substrate specificity of single and double mutants.

Through the overlay of the crystal structure of NAAR on that of MR, we determined that three residues (F52, E317, and L319) need to be mutated to widen the substrate-binding pocket to provide space for N-acetyl phenylglycine. To determine the important mutations, we first evaluated the single mutants.

i). F52V mutation decreased reactivity significantly: the variant only accepted (*R*)- and (*S*)-mandelate, both at similar rates but drastically reduced catalytic proficiency k_{cat}/K_M (10^5-fold) over the wild-type. In comparison, the Bearne group found a 12-fold decreased efficiency for the variant F52W as well as both decreased affinity for and turnover of (*S*)-mandelate with added steric bulk (F52W/Y54W variant) (*35*). As the mutated residue 52 is located in the hydrophobic (*S*)-pocket, we would expect F52V-MR to have decreased affinity for (*S*)-mandelate but

unaltered affinity for (*R*)-mandelate. However, we found increased K_M values and decreased k_{cat} values, similarly for both enantiomers. Therefore, the effect of the mutation is not limited to the (*S*)-pocket.

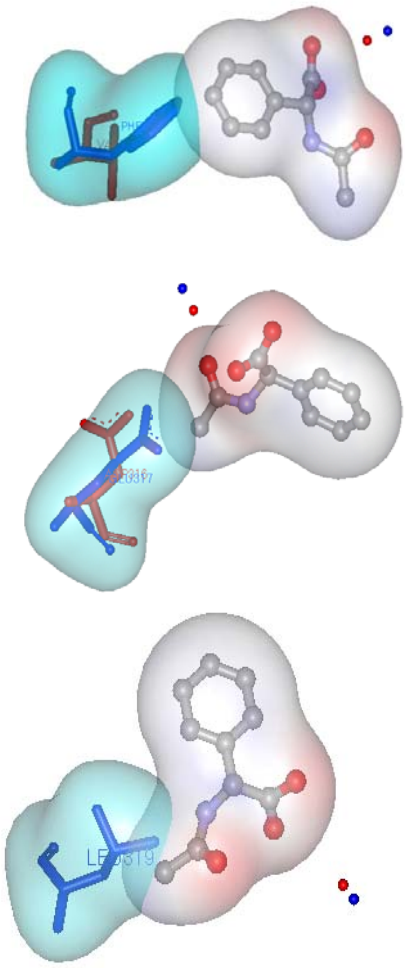
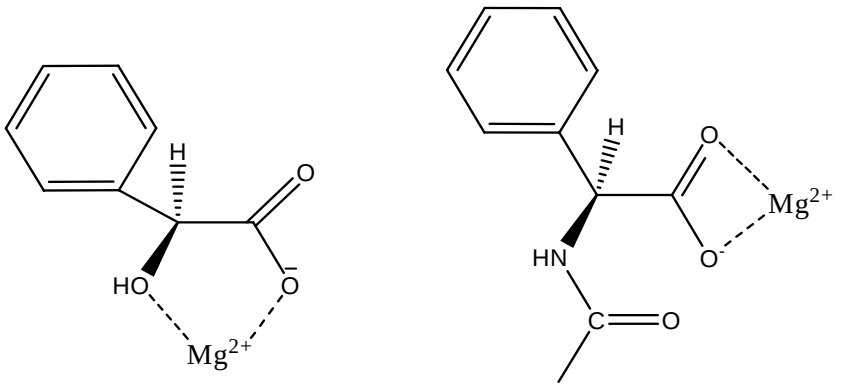


Figure 6-7. Residues that clash with N-acetyl-phenylglycine in substrate-binding pocket of MR template. A). F52 clashes, propose mutation to Val; B). E317 clashes, propose mutation to Asp; C). L319 clashes, propose mutation to Ala.



Mandelate N-acetyl-phenylglycine

Figure 6-8. Interaction of divalent metal ion with substrate for both mandelate, in MR active site (from 1MDR), and N-acetyl-phenylglycine in NAAR active site (from 1SJD).

ii). Since E317 is a catalytic residue, we would expect low, if any, activity for mandelate catalysis in a variant at that position. Indeed, in E317Q, Mitra et. al. (1994) found 3000-fold decreased catalytic efficiency for both enantiomers of mandelate and 4500-fold or 29000-fold decreased k_{cat} value for (R)- and (S)-mandelate, respectively. However, we found that the E317D single mutant, while featuring the expected low activity (four orders of magnitude reduced proficiency compared to MR, NAAR, or NSAR wild-type proteins), unexpectedly showed broadened specificity with low activity on nearly all of the tested substrates. The E317D mutation alone might be crucial for cross-reactivity.

iii). Despite any previous work on the importance of the residue L319 for catalytic activity of MR, we observed very little activity with mandelate and no activity with any other tested substrate on the L319A variant (Table 1, 4th column). Therefore, decreasing the size of residue 319 results in loss

of mandelate specificity but not in binding of any larger substrate. The capping domain of enolase superfamily members is known to be critical for substrate specificity but the environment around residue 319 in that domain differs significantly between NAAR and MR (*11*).

Next, we measured activity of the double mutants F52V/E317D, F52V/L319A, E317D/L319A with each of the substrates. All of the double mutants were inactive on all substrates except for both enantiomers of mandelate. However, the triple variant F52V/E317D/L319A shows activity on both enantiomers of mandelate, (*R*)-phenylglycine, N-methyl-(*R*)-phenylglycine, formyl-(*R*)-phenylglycine, N-acetyl-(*R*)-phenylglycine, and N-acetyl-phenylalanine (Table 6-3). None of the variants showed activity on N-acetyl-(*R*)-methionine. As hypothesized, an increase in the size of the substrate binding site enables cross-reactivity. We surmise that the orientation of the substrate is changed to accommodate the larger substrates. Also, interestingly this variant shows similar levels of catalytic specificity (k_{cat}/K_M = 780 M-1 s-1, on N-Ac-(*R*)-phenylglycine) to the OSBS/NAAR from *Amycolatopsis* sp. (k_{cat}/K_M = 370 M-1 s-1, on N-Ac-(*R*)-methionine) (*28*).

Table 6-3. Catalytic proficiencies for cross-reactivity mutations.

Substrate	k_{cat}/K_M (M^{-1} s^{-1})							
	MR	F52V	E317D	L319A	F52V/E317D	F52V/L319A	E317D/L319A	F52V/E317D/L319A
D-Mandelate	(1.9 ± 0.4) x 10^4	0.6 ± 0.08	8 ± 0.3	0.04 ± 0.0003	-	5 ± 0.6	9 ± 0.7	(6.6 ± 0.3) x 10^3
L-Mandelate	(1.3 ± 0.7) x 10^4	0.5 ± 0.03	10 ± 7	0.94 ± 0.003	13 ± 3	3 ± 0.5	5 ± 0.4	(4.5 ± 1.1) x 10^3
D-Phenylglycine	(2.2 ± 0.5) x 10^2	-	20 ± 0.5	-	-	-	-	(7.6 ± 2.4) x 10^3
L-Phenylglycine	(2.5 ± 0.5) x 10^2	-	20 ± 3	-	-	-	-	(4.8 ± 1.4) x 10^3
N-Methyl-D-Phenyglycine	-	-	8 ± 0.09	-	-	-	-	3 ± 0.3
N-Formyl-D-Phenyglycine	-	-	20 ± 0.5	-	-	-	-	(8.8 ± 3.2) x 10^3
N-Formyl-L-Phenyglycine	-	-	-	-	-	-	-	(9.9 ± 2.5) x 10^3
N-Acetyl-D-Phenylglycine	-	-	7 ± 0.2	-	-	-	-	(7.8 ± 1.3) x 10^2
N-Succinyl-D-Phenylglycine	-	-	2 ± 0.2	-	-	-	-	-
N-Acetyl-D-Phenylalanine	-	-	2 ± 0.2	-	-	-	-	13 ± 0.9

Table 6-4. Kinetic constants for F52V/E317D/L319A variant.

Substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
D-Mandelate	190 ± 40	29 ± 6	(6.6 ± 0.3) * 10^3
L-Mandelate	177 ± 37	39 ± 4	(4.5 ± 1.1) * 10^3
D-Phenylglycine	38 ± 5	5 ± 2	(7.6 ± 2.4) * 10^3
L-Phenylglycine	29 ± 5	6 ± 3	(4.8 ± 1.4) * 10^3
N-Formyl-D-Phenyglycine	291 ± 43	33 ± 3	(8.8 ± 3.2) * 10^3
N-Formyl-L-Phenyglycine	288 ± 60	29 ± 3	(9.9 ± 2.5) * 10^3
N-Methyl-D-Phenyglycine	0.56 ± 0.2	169 ± 33	3 ± 0.3
N-Acetyl-D-Phenylglycine	0.78 ± 0.09	143 ± 32	780 ± 130
N-Succinyl-D-Phenylglycine	-	-	-
N-Acetyl-D-Phenylalanine	0.47 ± 0.09	34 ± 5	13 ± 0.9
N-Acetyl-D-Methionine	-	-	-

6.4.3.2 Effects of the E317 mutations

Since residue E317 is deemed necessary for the stabilization of the enolate intermediate (*34*), we investigated the effects of charge and size on this residue by combining the substitutions E317D, D317N, D317E, D317Q, and D317A with the double mutant F52V/L319A (Table 6-5). We find that the catalytic efficiency of the F52V/D317**N**/L319A variant on both enantiomers of mandelate is similar to that of the F52V/E317**D**/L319A variant, and even slightly higher on N-acetyl-(*R*)-phenylglycine. Exchanging for the larger uncharged glutamine residue, however, resulted in much decreased activity on both mandelate enantiomers and no activity on N-acetyl-(*R*)-phenylglycine. However, both triple mutants with smaller residues in position 317 and thus more space for substrate, F52V/E317**A**/L319A and F52V/E317**D**/L319A, are proficient on both the

mandelate and N-acetyl-(*R*)-phenylglycine substrates again, the latter about ten times more than the former. From these results, we conclude that the size of the pocket near residue 317 is very important for high activity of MR but that the charge of residue 317 only plays a minor role. As E317 stabilizes the enolic intermediate through a short, strong hydrogen bond between the hydroxyl group of the enol and the anionic carboxylate group of E317 (*34*), E317D in all likelihood is not positioned for such a bond, and the uncharged residue in E317A might not bind the enolic intermediate at all.

6.4.3.3 Reduction of clash from N-succinyl-phenyglycine

Recently, N-succinyl-phenylglycine was found (*36*) to be an even better substrate for NAAR from *Amycolatopsis* than the early favorites, N-acetyl amino acids (*37*, *38*). NAAR is identical to OSBS and N-succinyl-phenylglycine mimics 2-succinyl-6R-hydroxy-2,4-cyclohexadiene-1R-carboxylate (SHCHC), the substrate of OSBS. When N-succinyl-phenylglycine is modeled into the active site of the triple variant (F52V/E317D/L319A) of MR, it clashes with the L319A mutation (Figure 6), even though the triple variant increases the size of the substrate-binding pocket. Indeed, as expected (as there is a clash with the substrate in the structure), the triple variant is inactive on N-succinyl-phenylglycine. A further increase of the substrate-binding pocket by mutating Leu319 to glycine instead of alanine yielded two contrary effects (Table 6-6): On

one hand, the triple variant F52V/E317D/L319**G** now accepts N-succinyl-phenylglycine; on the other hand, the new variant was now inactive on both enantiomers of phenylglycine and N-formyl-phenylglycine and showed four- to twenty-fold reduced activity on mandelate, N-methyl- and N-acetyl-phenylglycine as well as N-acetyl-phenylalanine. Remarkably, the single variant E317D is the only variant found in this work besides the new triple variant F52V/E317D/L319**G** that accepts N-succinyl-phenylglycine as a substrate; both double variants F52V/E317D and E317D/L319A as well as F52V/E317D/L319**A** are inactive. These findings demonstrate i) the non-additivity of effects of change of substrate specificity, and ii) the importance of the capping domain in the determination of the substrate specificity.

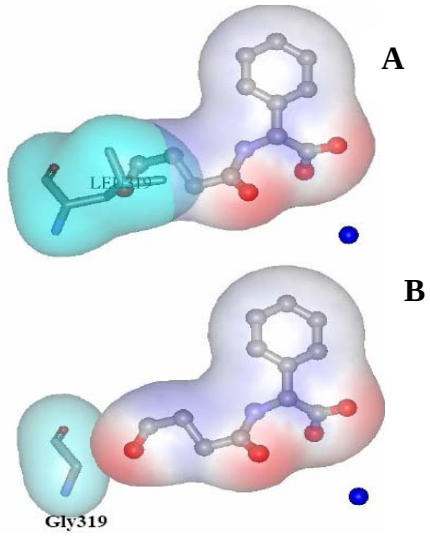


Figure 6-9. A). L319 clashes with N-succinyl-phenylglycine; B). Proposed mutation to Gly.

Table 6-5. Kinetic constants for F52V/E317X/L319A variants. T = 30 °C, pH 8.0, [E] = 0.3 μM.

Mutant	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
	D-Mandelate		
F52V/E317**D**/L319A	190 ± 40	29 ± 6	(6.6 ± 0.3) * 10^3
F52V/E317**N**/L319A	84 ± 8	54 ± 4	(1.6 ± 0.3) * 10^3
F52V/E317**E**/L319A	0.4 ± 0.05	88 ± 3	5 ± 0.6
F52V/E317**Q**/L319A	0.3 ± 0.01	56 ± 1	5 ± 0.6
F52V/E317**A**/L319A	48 ± 13	2 ± 0.2	(3.2 ± 5.5) * 10^4
	L-Mandelate		
F52V/E317**D**/L319A	177 ± 37	39 ± 4	(4.5 ± 1.1) * 10^3
F52V/E317**N**/L319A	84 ± 7	41 ± 1	(2.0 ± 0.3) * 10^3
F52V/E317**E**/L319A	0.2 ± 0.04	92 ± 8	2 ± 0.2
F52V/E317**Q**/L319A	0.98 ± 0.03	143 ± 11	7 ± 0.2
F52V/E317**A**/L319A	0.92 ± 0.02	4 ± 0.1	220 ± 2
	N-acetyl-D-phenylglycine		
F52V/E317**D**/L319A	0.78 ± 0.09	143 ± 32	780 ± 130
F52V/E317**N**/L319A	8 ± 2	51 ± 16	150 ± 9
F52V/E317**E**/L319A	-	-	-
F52V/E317**Q**/L319A	-	-	-
F52V/E317**A**/L319A	1 ± 0.1	3 ± 0.3	430 ± 2

Table 6-6. Kinetic constants for F52V/E317D/L319G variant. T = 30 °C, pH 8.0, [E] = 0.3 μM

Substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
D-Mandelate	10 ± 3	271 ± 110	1.9 ± 0.2
L-Mandelate	45 ± 4	1.2 ± 0.01	39 ± 0.4
D-Phenylglycine	-	-	-
L-Phenylglycine	-	-	-
N-Formyl-D-Phenyglycine	-	-	-
N-Formyl-L-Phenyglycine	-	-	-
N-Methyl-D-Phenyglycine	0.2 ± 0.02	21 ± 1.3	0.007 ± 0.0007
N-Acetyl-D-Phenylglycine	0.1 ± 0.01	21 ± 1.5	0.006 ± 0.0003
N-Succinyl-D-Phenylglycine	0.04 ± 0.002	19 ± 1.2	0.002 ± 0.0001
N-Acetyl-D-Phenylalanine	-	-	-
N-Acetyl-D-Methionine	-	-	-

6.4.4 Flexibility (Plasticity) of the active site residues

We have identified a triple mutant (F52V/E317D/L319A) that shows cross-reactivity across the MR and MLE subgroups. This finding allowed us to hypothesize that the exchange of active site residues between MR and NAAR would provide further proof of divergent evolution. The most straightforward way to realize this goal is to change the non-identical catalytic residues of MR to those of NAAR. A phenylacetone monooxygenase has been converted to a phenylcyclohexanone

monooxygenase using this rationale (*39*). All members of the enolase superfamily except the MR subgroup have two Lys residues, or a Lys and an Arg, on opposite sides of the active site that participate in the two-base mediated 1,1-proton transfer reaction, Lys163 and Lys263 in NAAR. These residues are located at the end of the second and sixth β-strands of the barrel, respectively. When overlaying the NAAR structure with that of MR, the general acid Lys166 is found at the end of the second β-strand of the barrel domain, the same location as Lys163 in NAAR. Lys263 in NAAR, however, superimposes with Asp270 in MR, part of a dyad of His297-Asp270 acting as a catalytic base. As D270 in MR and K263 in NAAR superimpose (*40*), we chose to mutate D270 to Lys, its structural homolog. For the same reason, we chose to mutate H297 to Gly. As the triple variant with the broadest substrate specificity, F52V/E317D/L319A, included N-acetyl-phenylglycine, a substrate of the MLE subgroup, we proceeded to add the two mutations suggested by the change of non-identical catalytic residues of MR to those of NAAR, D270K and H297G, to that triple variant F52V/E317D/L319A.

We found that the quintuple variant (F52V/E317D/L319A/D270K/H297G) was active on all substrates except N-acetyl-(*R*)-phenylalanine and N-succinyl-(*R*)-phenylglycine but including N-acetyl-(*R*)-phenylglycine (Table 6-7). To reveal any potential evolutionary path, we separately combined the catalytic residue

mutations D270K and H297G into the triple variant template (F52V/E317D/L319A). The finding that the latter variant, F52V/E317D/L319A/**D270K**, featured slight activity on both enantiomers of mandelate and phenylglycine, whereas the former variant F52V/E317D/L319A/**H297G** only shows activity on (*S*)-mandelate, provides evidence for the D270K residue playing the role of the catalytic residue in the latter variant.

Table 6-7. Kinetic constants for F52V/E317D/L319A/H297G/D270K mutant.

Substrate	k_{cat} (s^{-1})	K_M (mM)	k_{cat}/K_M (M^{-1} s^{-1})
D-Mandelate	0.01 ± 0.001	15 ± 2	0.7 ± 0.07
L-Mandelate	0.13 ± 0.02	22 ± 3	6 ± 0.7
D-Phenylglycine	0.02 ± 0.003	1.6 ± 0.6	10 ± 1
L-Phenylglycine	0.12 ± 0.04	11 ± 2.4	10 ± 2
N-Formyl-D-Phenyglycine	0.05 ± 0.004	0.9 ± 0.2	60 ± 7
N-Formyl-L-Phenyglycine	0.02 ± 0.001	7 ± 1	3 ± 0.3
N-Methyl-D-Phenyglycine	0.3 ± 0.2	2.6 ± 0.9	120 ± 20
N-Acetyl-D-Phenylglycine	0.08 ± 0.009	23 ± 2	3 ± 0.2
N-Succinyl-D-Phenylglycine	-	-	-
N-Acetyl-D-Phenylalanine	-	-	-

6.4.5 Implications for the evolution of the enolase superfamily.

Members of the three enolase superfamily subgroups are classified based on the acid/base functional groups located at the C-terminal end of the second, third, fifth, sixth, and seventh β-strands. The finding that conservation of these motifs has been observed in evolutionary trees generated from the overall sequences suggests that these motifs are important in understanding divergent evolution (*11*). The Gerlt group has done an exceptional job of providing evidence for divergent evolution in the MLE subgroup (*12*, *13*, *23*, *27*). However, the only evidence for divergent evolution across two subgroups of the enolase superfamily, until now, is the conserved bi-domain structure. Despite low sequence identity and low structural similarity among the MR and MLE subgroup capping domains we were able to show cross-reactivity <u>across</u> these two enolase superfamily groups through three mutations.

The three residues in the triple variant responsible for cross-reactivity across subgroups are located in the capping domain. Substitution of these residues relaxes the substrate specificity. The Gerlt group demonstrated this same phenomenon when its findings revealed cross-reactivity <u>within</u> the MLE subgroup through a single substitution in both AEE (D297G) and MLE II (E323G) to enable catalysis of the OSBS reaction. Both the substituted residues are located at the end of the eighth β-strand of the barrel domain, which leads into the capping domain (*23*). The same

group demonstrated that the difference between the structures of the OSBS from *Amycolatopsis sp*. (catalyzes both dehydration and racemization) and *E. coli* (catalyzes only dehydration, as of yet) is the loop following the β8 strand of the barrel domain, which provides more space for the N-succinylphenylglycine in the OSBS from *Amycolatopsis* and allows the OSBS to act promiscuously as an NAAR (*29*). Both the published and the presented results provide further proof that the capping domain, especially the loop following the β8-strand is responsible for substrate specificity.

Structural conservation of the $(\beta/\alpha)_7$ β-barrel domain provides a stable structural environment for the active site residues, with enough flexibility in the loops at the end of the β-strands to allow the evolution of different active sites. In the enolase superfamily the catalytic acid/base residues are mostly conserved, with a conserved catalytic base at the end of the β6-strand for all of the subgroups except for the MR subgroup. As we have shown through just one or three mutations (Table 6-3), E317D and F52V/E317D/L319A, we are able to produce a variant with promiscuous NAAR and MR activity. This has also been observed by the Gerlt group, who states that the divergent evolution of function in superfamilies where the active site structure "hard-wires" chemistry begins with relaxation of substrate specificity, which generates functional promiscuity (*13*, *23*).

6.4.6 Divergent evolution models.

There are several models that attempt to describe the path that enzymes take as they diverge. *Ohno's model* assumes that i) gene duplication is the first step to divergence and that ii) new function develops at the expense of the original function, as was shown for the presence of strong negative pressure in the evolution of β-lactamase (*3*, *19*). The alternate *gene-sharing model* assumes that first a generalist enzyme evolves which performs multiple metabolic functions, followed by gene duplication and divergent specialization (*41*). Often during directed evolution experiments generalist enzymes are evolved supporting this model of divergent evolution (*17*). Recently, Tawfik's group developed a model that comprises both Ohno's model, where gene duplication happens prior to speciation, and the gene-sharing model, where gene duplication occurs after speciation (*5*).

The enolase superfamily demonstrates in several published examples that both scenarios do occur. For example, strong negative trade-offs are observed for the evolution of OSBS activity from the AEE scaffold, with greater than 2400-fold increase of the new function and 7800-fold decrease of the original function (*23*). However, weak negative trade-offs are observed in the evolution of OSBS activity in the MLE II scaffold, with over a one million-fold increase in new function, and the old function decreasing by only about 15-fold (*23*). A natural occurring

generalist (weak trade-off) is observed with the OSBS/NSAR gene found by Sakai et al. 2006 with catalysis for the dehydration (OSBS reaction) with a (k_{cat}/K_M = 3.7 x 10^2 M^{-1} s^{-1}) and the racemization (NSAR reaction) with a (k_{cat}/K_M = 3.7 x 10^4 M^{-1} s^{-1}) for N-Succinyl-L-Ala (*12*). A strong negative trade-off is observed in the NAAR/OSBS from *Amycolatopsis* which catalyzes dehydration (OSBS reaction) efficiently (k_{cat}/K_M = 2.5 x 10^4 M^{-1} s^{-1}) but racemization (NAAR reaction) only very inefficiently (k_{cat}/K_M = 3.7 x 10^2 M^{-1} s^{-1}) (*28*). Our results and results from literature provide proof of divergent evolution (Figure 6-10). Results presented here show strong negative trade-offs with an increase in new function and a decrease in original function. A strong negative trade-off is expected as mutations through rational design will have a stronger effect than mutations through directed evolution. These results are in agreement with the hypothesis proposed by the Tawfik group stating that the trade-off is different and can be quantified.

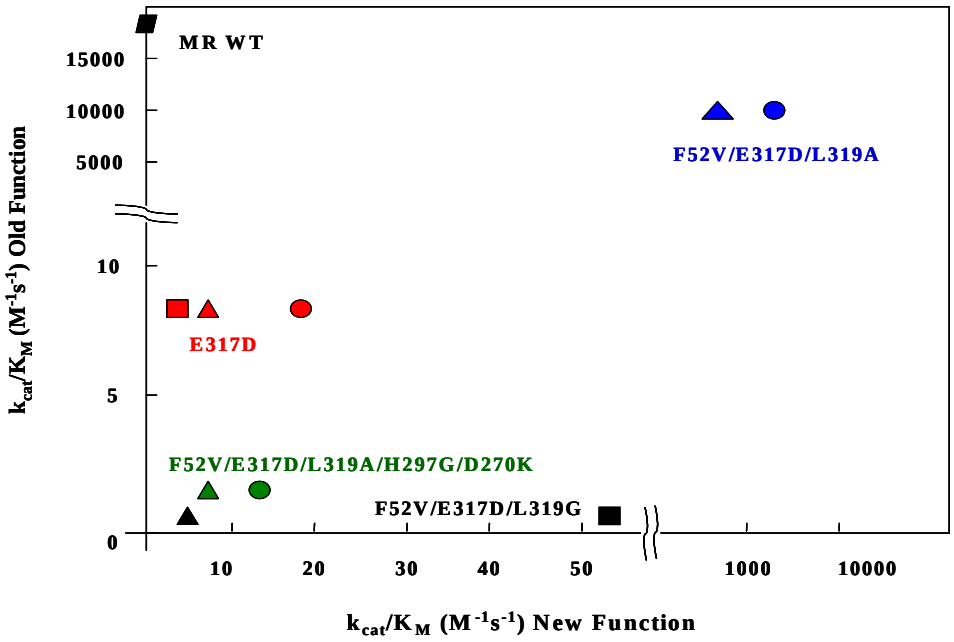


Figure 6-10. Map of variants in comparison to wild-type MR to reveal generalists and specialists. (○) Formyl-(*R*)-phenylglycine, (Δ) N-acetyl-(*R*)-phenylglycine, (□) N-succinyl-(*R*)-phenylalycine.

6.5 Conclusion

In our quest to investigate the scope of divergent evolution in the enolase superfamily, we employed mandelate racemase (MR) as a template and devised a group of site-directed mutations that broadened substrate specificity and redesigned the active site of MR toward that of N-acetylamino acid racemase (NAAR). We developed single, triple and quintuple mutants, each able to catalyze the racemization of substrates including mandelate, phenylglycine, N-methyl-phenylglycine, N-formyl-phenylglycine, N-acetyl-phenylglycine, and N-acetyl-phenylalanine. By just changing either a single amino acid, E317D, three amino acids,

F52V/E317D/L317A, we obtained variants able to racemize substrates typical for both MR (mandelate) and NAAR (N-acetyl-(*R*)-phenylglycine and –phenylalanine). For the first time, we have demonstrated possible divergent evolution across the MR and MLE subfamilies of the enolase superfamily; so far, only possible divergent evolution of function within the MLE and MR subfamilies was known. The E317D variant demonstrates strong trade-off according to the evolutionary model by Kheronsky et al.(*5*), as MR functionality (catalytic specificity, on (*R*)-mandelate) declines by more than three orders of magnitude, whereas NAAR functionality (on N-acetyl-(*R*)-phenylglycine) increases from immeasurably small levels to the same specificity as that of MR. There is no trade-off at all when changing to the triple variant F52V/E317D/L319A, as MR and NAAR specificities increase by three and two orders of magnitude, respectively. However, neither double variant shows much of any catalytic activity at all, and while both N-acetyl-(*R*)-phenylglycine and –phenylalanine are accepted as substrates, N-acetyl-(*R*)-methionine is not, demonstrating the crucial role of the phenyl group. We conclude that although a 'generalist' can be created within the MLE subfamily with just one amino acid substitution to enable catalysis of the OSBS reaction (D297G to change AEE to a promiscuous AEE/OSBS and E323G to change MLE II into a MLE II/OSBS), the creation of a 'generalist' for the entire enolase

superfamily, catalyzing MR and NAAR reactions, can be achieved just as well by a single mutation (E317D).

6.6 REFERENCES

(1) Darwin, C. (1859) *The Origin of Species*, Signet classic, New York, NY.

(2) Bridgham, J. T., Carroll, S. M., and Thornton, J. W. (2006) Evolution of Hormone-Receptor Complexity by Molecular Exploitation. *Science 312*, 97-101.

(3) Ohno, S. (1970) *Evolution by Gene Duplication*, Springer Verlag, New York.

(4) Jensen, R. A. (1976) Enzyme Recruitment in Evolution of New Function. *Annu. Rev. Microbiol 30*, 409-425.

(5) Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme Promiscuity: Evolutionary and Mechanistic Aspects. *Current Opinion in Chem. Biol. 10*, 498-508.

(6) O'brien, P. J., and Herschlag, D. (1999) Catalytic Promiscuity and the Evolution of New Enzymatic Activities. *Chem. Biol. 6*, R91-R105.

(7) Copley, R. R., and Bork, P. (2000) Homology among (Beta Alpha)(8) Barrels: Implications for the Evolution of Metabolic Pathways. *Journal of Molecular Biology 303*, 627-640.

(8) Lenski, R. E., Ofria, C., Pennock, R. T., and Adami, C. (2003) The Evolutionary Origin of Complex Features. *Nature 423*, 139-144.

(9) Meng, E. C., Polacco, B. J., and Babbitt, P. C. (2004) Superfamily Active Site Templates. *Proteins: Structure, Function, and Bioinformatics 55*, 962-976.

(10) Glasner, M. E., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of Enzyme Superfamilies. *Current Opinion in Chem. Biol. 10*, 492-497.

(11) Gerlt, J. A., Babbitt, P. C., and Rayment, I. (2005) Divergent Evolution in the Enolase Superfamily: The Interplay of Mechanism and Specificity. *Arch. of Biochem. and Biophys. 433*, 59-70.

(12) Sakai, A., Xiang, D. F., Xu, C., Song, L., Yew, W. S., Raushel, F. M., and Gerlt, J. A. (2006) Evolution of Enzymatic Activities in the Enolase Superfamily: N-Succinylaminoacid Racemase and a New Pathway for the Irreversible Conversion of D- to L-Amino Acids. *Biochemistry 45*, 4455-4462.

(13) Vick, J. E., Schmidt, D. M., and Gerlt, J. A. (2005) Evolutionary Potential of $(B/\alpha)8$-Barrels: In Vitro Enhancement of A "New" Reaction in the Enolase Superfamily. *Biochemistry 44*, 11722-11729.

(14) Hocker, B. (2005) Directed Evolution of $(B\alpha)_8$-Barrel Enzymes. *Biomolecular Engineering 22*, 31-38.

(15) Babbitt, P. C., Hasson, M. S., Wedekind, J. E., Palmer, D. R. J., Barrett, W. C., Reed, G. H., Rayment, I., Ringe, D., Kenyon, G. L., and Gerlt, J. A. (1996) The Enolase Superfamily: A General Strategy for Enzyme-Catalyzed Abstraction of the Alpha-Protons of Carboxylic Acids. *Biochemistry 35*, 16489-16501.

(16) Gerlt, J. A., and Babbitt, P. C. (2000) Can Sequence Determine Function? *Genome Biology 1*.

(17) Aharoni, A., Gaidukov, L., Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2005) The 'Evolvability' of Promiscuous Protein Functions. *Nat. Genet. 37*, 73-76.

(18) Collins, C. H., Leadbetter, J. R., and Arnold, F. H. (2006) Dual Selection Enhances the Signaling Specificity of a Variant of the Quorum-Sensing Transcriptional Activator Luxr. *Nat. Biotechnol. 24*, 708-712.

(19) Varadarajan, N., Gam, J., Olsen, M. J., Georgiou, G., and Iverson, B. L. (2005) Engineering of Protease Variants Exhibiting High Catalytic

Activity and Exquisite Substrate Selectiviey. *PRoc. Natl. Acad. Sci. 102*, 6855-6860.

(20) Farber, G. K., and Petsko, G. A. (1990) The Evolution of a/B Barrel Enzymes. *Trends Biochem Sci 15*, 228-234.

(21) Nagano, N., Hutchinson, E. G., and Thornton, J. M. (1999) Barrel Structures in Proteins: Automatic Identification and Classification Including a Sequence Analysis of Tim Barrels. *Protein Science 8*, 2072-2084.

(22) Gerlt, J. (2000) New Wine from Old Barrels. *Nat. Struct. Biol. 7*, 171-173.

(23) Schmidt, D. M. Z., Mundorff, E. C., Dojka, M., Bermudez, E., Ness, J. E., Govindarajan, S., Babbitt, P. C., Minshull, J., and Gerlt, J. A. (2003) Evolutionary Potential of (Beta/Alpha)(8)-Barrels: Functional Promiscuity Produced by Single Substitutions in the Enolase Superfamily. *Biochemistry 42*, 8387-8393.

(24) Babbitt, P. C., and Gerlt, J. A. (1997) Understanding Enzyme Superfamilies - Chemistry as the Fundamental Determinant in the Evolution of New Catalytic Activities. *Journal of Biological Chemistry 272*, 30591-30594.

(25) Neidhart, D. J., Kenyon, G. L., Gerlt, J. A., and Petsko, G. A. (1990) Mandelate Racemase and Muconate Lactonizing Enzyme and Mechanistically Distinct and Structurally Homologous. *Nature 347*, 692-694.

(26) Lumry, R., and Eyring, H. (1954) Conformation Changes of Proteins. *J Phys Chem 58*, 110-120.

(27) Glasner, M. E., Fayazmanesh, N., Chiang, R. A., Sakai, A., Jacobson, M. P., Gerlt, J. A., and Babbitt, P. C. (2006) Evolution of Structure and Function in the O-Succinylbenzoate Synthase/N-Acylamino Acid Racemase Family of the Enolase Superfamily. *J. Mol. Biol. 360*, 228-250.

(28) Palmer, D. R. J., Garrett, J. B., and Gerit, J. A. (1999) Unexpected Divergence of Enzyme Function and Sequence: "N- Acylamino Acid Racemase" From Amycolaptosis Is an Ortho- Succinylbenzoate Synthase. *Faseb Journal 13*, A1365-A1365.

(29) Thoden, J. B., Taylor Ringia, E. A., Garrett, J. B., Gerlt, J. A., Holden, H. M., and Rayment, I. (2004) Evolution of Enzymatic Activity in the Enolase Superfamily: Structural Studies of the Promiscuous *O*-Succinylbenzoate Synthase from *Amycolatopsis*. *Biochemistry 43*, 5716-5727.

(30) Schmidt, D.M, Mitra, B. E., Doijka, M., Bermudez, E., Ness, J.E., Govindarajan, S., Babbitt, P.C., Minshull, J., and Gerlt, J.A. (2003) Evolutionary Potential of (B/a)8-Barrels: Functional Promiscuity Produced by Single Substitutions in the Enolase Superfamily. *Biochemistry 42*, 8387-8393.

(31) Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, Vol. 3, Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

(32) Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) The Protein Data Bank. *Nucleic Acids Res. 28*, 235-242.

(33) Felfer, U., Goriup, M., Koegl, M. E., Wagner, U., Larissegger-Schnell, B., Faber, K., and Kroutil, W. (2005) The Substrate Spectrum of Mandelate Racemase: Minimum Structural Requirements for Substrates and Substrate Model. *Advanced Synthesis & Catalysis 347*, 951-961.

(34) Mitra, B., Kallarakal, A. T., Kozarich, J. W., Gerlt, J. A., Clifton, J. G., Petsko, G. A., and Kenyon, G. L. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Importance of Electrophilic Catalysis by Glutamic Acid 317. *Biochemistry 34*, 2777-2789.

(35) Siddiqi, F., Bourque, J. R., Jiang, H., Gardner, M., St. Maurice, M., C., B., and S.L., B. (2005) Perturbing the Hydrophobic Pocket of

Mandelate Racemase to Probe Phenyl Motion During Catalysis. *Biochemistry 44*, 9013-9021.

(36) Taylor Ringia, E. A., Garrett, J. B., Thoden, J. B., Holden, H. M., Rayment, I., and Gerlt, J. A. (2004) Evolution of Enzymatic Activity in the Enolase Superfamily: Functional Studies of the Promiscuous O-Succinylbenzoate Synthate from *Amycolatopsis*. *Biochemistry 43*, 224-229.

(37) Tokuyama, S. (1998) Studies on Function and Application of N-Acylamino Acid Racemase. *Nippon Nogeikagaku Kaishi-J. Jpn. Soc. Biosci. Biotechol. Agrochem. 72*, 137-144.

(38) Verseck, S., Bommarius, A., and Kula, M. R. (2001) Screening, Overexpression and Characterization of an N- Acylamino Acid Racemase from Amycolatopsis Orientalis Subsp Lurida. *Appl. Microbiol. Biotechnol. 55*, 354-361.

(39) Bocola, M., Schulz, F., Leca, F., Vogel, A., Fraaije, M. W., and Reetz, M. T. (2005) Converting Phenylacetone Monooxygenase into Phenylcyclohexanone Monooxygenase by Rational Design: Towards Practical Baeyer-Villiger Monooxygenases. *Adv. Synth. Catal. 347*, 979-986.

(40) Thompson, T. B., Garrett, J. B., Taylor, E. A., Meganathan, R., Gerlt, J. A., and Rayment, I. (2000) Evolution of Enzymatic Activity in the Enolase Superfamily: Structure of O-Succinylbenzoate Synthase from *Escherichia Coli* in Complex with Mg^{2+} and O-Succinylbenzoate Synthase. *Biochemistry 39*, 10662-10676.

(41) Piatigorsky, J., and Wistow, G. J. (1989) Enzyme/Crystallins: Gene Sharing as an Evolutionary Strategy. *Cell 57*, 197-199.

CHAPTER 7

CONCLUSIONS AND FUTURE WORK

7.1 Conclusions

The search for extraterrestrial life has long been a subject of human interest. Technological advances in the last decades have allowed us to come closer to proving/disproving the existence of extraterrestrial life. Assuming that formation of life requires liquid water in addition to homochiral molecules and redox energy, there are locations of interest in our solar system, which include Mars, Enceledus, Europa, and Titan.

When searching for life we are interested in finding homochiral molecules as these are representative of life molecules. Although, some chemicals as inanimate catalysts in clays can catalyze chiral reactions, in most cases the change in chirality is related to a biological reaction. Therefore in the search for extraterrestrial life it is advantageous to search for a change in chirality; a racemization reaction is a good representation of a reaction displaying such chirality change.

The goal of this project was to develop a benchmark to be used in the search for extraterrestrial life. Through each of the four Aims reported

in this work we have developed such a benchmark with mandelate racemase.

The first aim involved the cloning of the mandelate racemase gene, along with the overexpression, and characterization of the protein from an *Escherichia coli* host. We were able to clone the gene from the *Pseudomonas putida* genome, overexpress it in the vector pkk223, and characterize it. Kinetic data for the mandelate racemase reaction gained with the polarimetric assay employed in this work, however, differed significantly from several sets of data reported previously, obtained by other assays such as chromatography and hydrogen-deuterium exchange. The current shortcomings of polarimetry with respect to response time and sensitivity might be responsible for this finding. We conclude that further work to improve polarimetry along both dimensions of merit is urgently required.

The second aim was to investigate the activity of mandelate racemase at low temperatures and in low water media. We found the enzyme was not active in organic solvents at the concentrations needed for low temperature measurements. These results are similar to what the Faber group found where mandelate racemase is inactive in organic solvents, but not deactivated (*1*), most likely due to the need for high water activity for catalytic activity. Mandelate racemase was active in both concentrated salt solutions and water-in-oil microemulsions in a

temperature range of –30°C to 70°C. We found that the Gibbs free enthalpy of reaction for the water-in-oil microemulsions increased by more than half from both the aqueous and concentrated ammonium salt media. Thus, we have demonstrated that both concentrated ammonium salt solutions and water-in-oil microemulsions, both believed to be able to form on extraterrestrial planets and moons in the presence of liquid water, are suitable media for enzyme reactions at subzero temperatures.

Through exploration into the thermal stability of mandelate racemase we found that its unfolding is irreversible. We investigated three models: a two-state model, consecutive two-step irreversible model, and the Lumry-Eyring model that could represent the mechanism for this irreversible unfolding. This type of modeling has not been done before for an enzyme of this size and multimericity. Analyzing thermostability of large enzymes is difficult because in most cases, after the subunits unfold and the hydrophobic regions are exposed the subunits tend to aggregrate which does not allow for further measurements via DSC. We found the enzyme does not associate during thermal denaturation and the experimental data best fit the two-state irreversible model.

When searching for extraterrestrial life, insight into the natural evolution of enzymes with α/β-barrel structure is of interest, since this is the most abundant structure found in all enzymes, exploration into this was the fourth aim. We chose to investigate a common ancestor for the enolase

superfamily, by providing proof for divergent evolution between two of the subgroups. We overlayed the structure of MR, from the MR subgroup, with the structure of NAAR, from the MLE subgroup and identified three residues in the capping domain of MR that clashed with the NAAR substrate. We found both a single and triple mutant that were able to catalyze the racemization of mandelate, N-acetyl-phenylglycine, and N-succinyl-phenylglycine, showing cross-reactivity between the MR and MLE subgroups. For further proof of divergent evolution between these subgroups we also investigated the plasticity of the active site. We discovered a quintuple mutant that again catalyzed substrates from both the MR and MLE subgroups. Therefore, through these variants we were able to create generalist for the enolase superfamily, according to the Tawfik model (*2*).

Lastly, we conclude that reactions yielding a polarimetric signal, such as racemizations employed in this work, are very suitable as a tool to utilize chirality to find signs of life by observing the change of chirality with time rather than just a static chiral signal. We have developed the enzyme mandelate racemase as a benchmark in the search for extraterrestrial life.

7.2 Future Work

Mandelate racemase makes a good test case in the search for extraterrestrial life. The most important improvement necessary for this

work is the development of a more sensitive polarimeter for the enzymatic assay. Improved sensitivity in the polarimeter will allow for more accurate data, superior to anything that has been previously done. Immediate application of such a novel polarimeter is measuring kinetic data of the enzyme. We found a K_M value of about 10 mM, while those reported in literature using other assays are less (about 0.4 mM (*3*)). Improved sensitivity of the instrument will also have other benefits, including measurement of the enzymatic reaction at subzero temperatures. There seems to be slight activity in the 60% saturated concentrated salt solutions, however we were unable to make feasible measurements in the current polarimetry setup. Detecting activity in higher concentrations of ammonium salts will allow for measurement at even lower temperatures than those already found. Besides sensitivity of the instrument, another polarimeter modifications would be a smaller sample volume, currently the sample volume is large (8 mL), a smaller sample cell would allow for the need of less substrate and enzyme.

Enzymatic activity at low temperatures could also be enhanced by engineering the protein to be more flexible, since it has been determined that psychrophilic enzymes are more flexible than their mesophilic counterparts (*4*). Some reasons for this flexibility include: reduced core hydrophobicity, decreased ionic and electrostatic interactions, additional surface loops, fewer interdomain, subunit, and aromatic stacking

interactions (*4*). Psychrophilic enzymes also have lower proline content, lower number of disulfide bonds as well as more glycine clusters (*5*). Knowing this it would be possible to make mutations within the MR gene that will allow for increased activity at lower temperatures than that previously found.

Another topic of interest is to further explore the unfolding of the enzyme. Further studies could determine how to dissociate the enzyme and then to determine if the monomer, dimer, and tetramer are active. Investigation of this is important for two reasons. First, MR is an ancestral enzyme and it is would be of scientific significance to determine the necessity of the octameric structure. Second, it would be more efficient to work with a monomeric structure as opposed to the more complex octamer.

Towards this goal, it is important to determine the specific amino acids involved in the protein-protein binding interactions. From here we can try to mutate these amino acids in a manner to generate enough repulsive interactions to result in a monomeric form of the enzyme.

7.3 REFERENCES

(1) Felfer, U., Goriup, M., Koegl, M.F., Wagner, U., Larissegger-Schnell, B., Faber, K., and Kroutil, W. (2005) The Substrate Spectrum of Mandelate Racemase: Minimum Structural Requirements for Substrates and Substrate Model. *Adv. Synth. Catal. 347*.

(2) Khersonsky, O., Roodveldt, C., and Tawfik, D. S. (2006) Enzyme Promiscuity: Evolutionary and Mechanistic Aspects. *Current Opinion in Chem. Biol. 10*, 498-508.

(3) Kallarakal, A.T., Mitra, B., Kozarich, J.W., Gerlt, J.A, Clifton, J.G., Petsko, G.A., Kenyon, G.l. (1995) Mechanism of the Reaction Catalyzed by Mandelate Racemase: Structure and Mechanistic Properties of the K166r Mutant. *Biochemistry 34*, 2788-2797.

(4) Georlette, D., Blaise, V., Collins, T., D'amico, S., Gratia, E., Hoyoux, A., Marx, J. C., Sonan, G., Feller, G., and Gerday, C. (2004) Some Like It Cold: Biocatalysis at Low Temperatures. *Fems Microbiology Reviews 28*, 25-42.

(5) Gerday, C., Aittaleb, M., Arpigny, J. L., Baise, E., Chessa, J. P., Garsoux, G., Petrescu, I., and Feller, G. (1997) Psychrophilic Enzymes: A Thermodynamic Challenge. *Biochimica et Biophysica Acta 1342*, 119-131.